词汇获得年龄效应的认知机制：三水平元分析*

张露† 王子谦† 张清芳

（中国人民大学心理学系，北京 100872）

摘 要 词汇获得年龄(Age of acquisition, AoA)指个体最早理解某个单词意义时的年龄，研究发现早获得词汇比晚获得词汇的加工速度更快，准确率更高，表现出词汇的 AoA 效应。针对 AoA 效应的认知机制，研究者提出了多层次理论、表征理论和任意映射理论。为了系统地考察 AoA 效应发生的认知机制，本研究使用三水平元分析对考察 AoA 效应的已有文献进行定量分析，研究中检索和筛选了 2024 年 5 月 15 日以前发表的文献，最终纳入 130 篇研究(271 个效应量，N = 10697)。元分析结果发现 AoA 效应的效应量大小为 Hedge' s $g = 0.47$($p < 0.001$, 95%CI = [0.41, 0.52])，其效应大小受任务类型和词频的调节，并且主要依赖于不同任务的语义相关度，但不受目标认知过程、文字系统和任务语音相关度的影响。研究结果表明 AoA 主要影响了语义加工阶段，支持了表征理论，早期习得的词汇具有更丰富的语义网络，从而促进了词汇加工过程。

关键词 图画命名，词汇朗读，词汇判断，词汇获得年龄效应，三水平元分析

分类号 B842

收稿日期：2024-10-07

*国家自然科学基金面上项目(32171055)和国家社科基金重大项目（24&ZD251）资助。

† 张露和王子谦为共同第一作者。

通信作者：张清芳，E-mail: qingfang.zhang@ruc.edu.cn

1 引言

词汇获得年龄(Age of acquisition, AoA)是指个体在语言获得与发展过程中，最早接触到某个词并理解其意义的年龄。研究者发现在词汇理解(Rochford & Williams, 1962a, 1962b)和词汇产生(Carroll & White, 1973a)过程中，早获得词汇比晚获得词汇的加工速度更快，正确率更高，这一现象被称作 AoA 效应。AoA 效应具有跨语言的普遍性，研究者在英语、西班牙语、荷兰语、汉语等多种语言中均观察到了这一效应(Barry et al., 2006; Cuetos et al., 1999; Ellis & Morrison, 1998; 陈宝国 等, 2004)。以往研究者对于 AoA 效应的认知机制存在争论，因此，我们使用三水平元分析的方法定量分析以往涉及 AoA 效应的行为研究，并纳入调节变量探索 AoA 效应的认知机制。

1.1 AoA 效应的认知机制

研究者主要争论的是 AoA 效应发生在词汇加工过程中的哪个阶段。目前，针对 AoA 效应认知机制的主流观点包括多层次理论(The Multiple Loci Account)、表征理论(Representation Theory)和任意映射假设(Arbitrary Mapping Hypothesis)，研究者在表征理论与任意映射理论的基础上提出了整合理论(The Integrated View of AoA)。考察 AoA 效应的加工任务根据刺激材料分为两类：一类是以词汇作为刺激材料的任务，包括词汇判断(Johnston & Barry, 2006)、词汇朗读(Barry et al., 2006)和语义范畴判断(Chen et al., 2007)等，另一类是以图画作为刺激材料的任务，包括图画语义分类(Catling & Johnston, 2006a)、图画知觉辨认任务(Holmes & Ellis, 2006)、客体图画口语命名(Catling et al., 2008; Laganaro & Perret, 2011; Perret et al., 2014)和动作图画口语命名(Bonin, Boyer, et al., 2004; Shao et al., 2014)等。人们对词汇和图画两类材料的加工过程是不同的。词汇加工是从字形通达到语义的加工过程，而图画加工是从概念激活开始的加工过程，在词汇产生任务中包括了概念准备、词汇选择(基于概念的激活选择相应的词汇，包括句法和语义的选择)、音韵编码(提取音段信息和节律信息)、语音编码(编码各类信息的发音运动程序)和发音等过程(Levelt et al., 1999)。

多层次理论认为，AoA 效应来源于形、音、义加工的不同层面或者阶段。具体而言，AoA 并非单独地影响词汇加工或图画加工中的语音提取或语义提取等某一特定阶段，而是对不同加工阶段均会产生影响(Moore et al., 2004)。多层次理论是建立在早期的语音完整性假设基础上的，语音完整性假设(Phonological Completeness Hypothesis)认为 AoA 效应仅发生在词汇加工的语音提取阶段，早获得词语音以整体形式存储，而晚获得词语语音以非整体分解的形式存储，包括各个音段以及节律信息等，因此早获得词汇的提取速度快于晚获得词

(Brown & Watson, 1987)。研究者在真假词的词汇判断任务中操纵语音相关的变量，发现了AoA 效应；在词汇判断的同时完成发音抑制任务，无法进行语音提取时，AoA 效应则消失(Gerhand & Barry, 1999b)。也有研究发现 AoA 效应并不依赖整体或部分语音的提取(张振军等, 2011)，反对语音完整性假设的观点。后续研究发现 AoA 效应也出现在反映感知觉加工阶段的图画真实性判断任务中(Moore et al., 2004)，据此 Moore 等人提出了多层次理论。在图画真实性判断任务中同时进行发音抑制后仍然发现了较小但达到显著水平的 AoA 效应(Holmes & Ellis, 2006)，为多层次理论提供了支持证据。Catling 和 Johnston(2006a)在多层次理论基础上提出了累积假说(Accumulation Hypothesis)，认为 AoA 会影响词汇加工的多个阶段，包括感知觉加工、概念加工、语义加工和音韵加工等。研究发现随着实验任务中所包括的加工阶段增多，AoA 效应增大。通过比较图画-图画判断任务、语义范畴判断(物体真实性判断)任务、词汇-图画判断任务和图画命名任务，研究者发现在图画命名任务中 AoA 效应量最大(Catling & Johnston, 2009)，这是由于图画命名任务中包括了上述所有 AoA 效应可能出现的加工阶段。采用高时间分辨率的脑电测量，研究者也发现 AoA 影响了图画命名中的语义和音韵编码等多个加工阶段(Laganaro & Perret, 2011; Perret et al., 2014; Valente et al., 2014)，为多层次理论和累积假说提供了支持证据。

表征理论的提出基于 AoA 效应与词频效应之间的联系和区别。该理论认为 AoA 效应包括了两个成分：其中一个成分与词频密切相关，而另一个成分则独立于词频(Brysbaert & Ghyselinck, 2006)。研究发现在词汇判断、词汇朗读和语义范畴判断任务中，词频效应和 AoA 效应的大小类似。依赖于词频的这部分 AoA 效应可能作用于知觉或音韵的单词形式处理水平；而独立于词频的 AoA 效应则源自概念和(或)语义表征网络内部的竞争，当某一特定概念(例如，水果)对应多个词汇(苹果、榴莲)时，与晚获得词汇相比，早获得词汇能作为这一概念范畴的典型实例(如苹果)，加工更容易(Brysbaert & Ghyselinck, 2006)。此外，表征理论认为早获得词汇具有更强的侧抑制，这会降低其他竞争项的激活水平，晚获得词汇需要克服来自早获得词汇的抑制作用才能得到激活(Belke et al., 2005)。以上过程的发生依赖于早晚获得词在语义网络中的特征差异，与晚获得词汇相比，早获得词汇的语义表征网络更加丰富(Steyvers & Tenenbaum, 2005)。词汇提取过程中，早获得词汇在语义网络中的节点联结程度更高，因此早获得词汇的语义网络更容易通达。研究者在语义联想任务中要求被试报告目标词出现时首先联想到的词汇，在语义分类任务中要求被试根据目标词的语义信息进行分类。研究者在这两类语义加工任务中均观察到了典型的 AoA 效应，表明 AoA 效应发生于语义加

工阶段(Brysbaert et al., 2000)；在多个涉及语义层面加工的判断任务中，如有无生命性判断任务(Morrison & Gibbons, 2006)、人造物体判断任务(Catling & Johnston, 2006c)，都发现了显著的 AoA 效应。在不包括语音提取过程的任务中(如面孔识别、语义分类等任务)，同样观察到了 AoA 效应，据此研究者提出了表征理论，认为 AoA 效应来自于语义系统中词汇语义表征的组织规则，由早获得词汇塑造的语义系统是晚获得词汇意义建立的基础(Brysbaert et al., 2000; Holmes & Ellis, 2006)。脑成像研究结果也表明早获得词在与语义记忆有关的楔前叶上有更强的激活，支持了早获得词拥有更丰富的语义联结网络的观点(Fiebach et al., 2003)。

任意映射理论的提出受到了平行分布式加工观点(Seidenberg & McClelland, 1989)的影响。根据任意映射理论，AoA 效应源自学习系统的特征，主要与表征水平(知觉、正字法、语义和语音等)之间的联结方式有关(Anderson & Cottrell, 2001; Ellis & Lambon Ralph, 2000; Monaghan & Ellis, 2010)，而不是位于某一个单独分离的系统(如语义系统)。根据联结主义神经网络的观点，研究者证实了人们在词汇获得与发展中将习得词汇纳入心理词典中时，早获得词汇的神经网络具有较高的可塑性，易于形成和修正输入和输出表征，随着心理词典的形成和发展，网络的可塑性下降。也就是说，早获得词汇对词汇联结网络的影响较大(Ellis & Lambon Ralph, 2000)。可塑性降低使得晚获得词汇更难在心理词典中加以巩固，导致晚获得词难以改变先前的网络结构，层级间也更难建立联结(Yum & Law, 2019)。

另一方面，如果晚获得词汇形素-音素的映射更为规则，在词汇加工中能够利用相似单词中正字法和语音之间的映射知识，这在一定程度上能弥补晚获得词汇所产生的加工劣势(Lambon Ralph & Ehsan, 2006; Zevin & Seidenberg, 2002)。当映射关系不规则时，早晚获得词之间就会表现出 AoA 效应，当映射关系规则时，AoA 效应较小，源于累积词频效应(Zevin & Seidenberg, 2004)。研究者采用词汇朗读任务考察了 AoA 与正字法-语音映射一致性的关系，发现二者间存在交互作用；与一致映射的条件相比，不一致映射条件下的 AoA 效应更大，表明在不一致映射条件下，晚获得词难以利用现有神经网络的映射关系，需要付出较大努力才能完成加工(Monaghan & Ellis, 2002a)。操纵汉字的词汇获得年龄与正字法-语义映射关系也得到了相似的结论(Chen et al., 2007)。ERP 研究中使用视觉词汇判断任务发现，当正字法与语音映射不一致时，晚获得词比早获得词引发更大的 LPC，影响了词汇加工的晚期过程，支持了映射理论的观点(Bakhtiar et al., 2016)。

表征理论与多层次理论强调了 AoA 对词汇加工过程的影响，任意映射理论强调了 AoA 对不同层次表征间映射程度的影响，这三种理论并不对立(Elsherif & Catling, 2022)。最近，

基于已有研究发现和各种理论观点，研究者提出了 AoA 效应的整合理论(Brysbaert & Ellis, 2016; Chang & Lee, 2020; Cortese et al., 2020)，认为早、晚获得词的表征差异以及表征间联结程度的差异共同导致了词汇加工效率的不同。AoA 会影响神经网络中表征的形成以及表征间的联结强度(Elsherif & Catling, 2022)，早获得词位于神经网络的中心，具有丰富的语义表征以及概念间联系，更适合早获得词的输入与输出；而晚获得词则难以与其他词建立联系，由于神经可塑性的降低，也变得难以改变已形成的神经网络，加工晚获得词需要付出更多努力以适应其特定的输入和输出关系。采用词汇朗读和词汇判断任务，研究者不仅观察到了 AoA 效应，而且发现 AoA 与词汇的正字法-语音一致性之间的交互作用共同影响了词汇朗读过程，表明词汇的表征和词汇表征之间的映射关系共同影响了词汇加工过程，支持了整合理论的观点(Chang et al., 2019; Chang & Lee, 2020)。

1.2 影响 AoA 效应的因素

采用不同的语言任务，研究者在不同语言中都观察到了 AoA 效应，已有发现表明影响 AoA 效应的重要因素包括了任务类型、文字系统、词频等多个因素(Elsherif et al., 2023)。

1.2.1 任务类型

根据目标认知过程可将词汇加工区分为词汇理解与词汇产生两类过程。词汇理解侧重于对输入的感知觉加工以及对词汇正字法、语音与语义的通达，词汇产生则强调在心理词典中选择合适的单词并进行语音编码与输出。探测词汇理解的任务主要包括词汇真假判断、语义范畴判断等，探测词汇产生的任务主要包括图画命名、词汇朗读等。词汇真假判断任务要求被试根据词汇类别区分呈现的单词是否为真词(González-Nosti et al., 2014; Izura & Hernández-Muñoz, 2017; Xu et al., 2021)，其中作为填充条件的实验材料是符合正字法与发音规则但无意义的假词(但不符合正字法规则的研究见 Gerhand & Barry, 1999b; Ghyselinck, Lewis, Brysbaert, 2004)，其加工过程的主要目的是提取单词的意义，包括了对词汇正字法、语义和(或)语音信息的提取。语义范畴判断任务中要求被试判断呈现的词汇是否属于某一指定的语义范畴(如人造物品)(Catling & Johnston, 2006c)，包括对词汇正字法和语义等信息的提取。图画命名任务一般要求被试对呈现的图画进行口语命名，是典型的词汇产生任务，包括了词汇语义编码和音韵编码等各个加工阶段。词汇朗读任务要求被试朗读所呈现的单词，主要包括了词汇产生过程中的音韵编码、语音编码和发音等过程，较少涉及对目标词语义的通达。研究者也采用了词汇补全任务(Gilhooly & Gilhooly, 1979)、模糊字辨认任务(陈宝国 等, 2010)和语音分割任务(Monaghan & Ellis, 2002a)等任务探索 AoA 效应发生的认知机制。

早期三类理论均认为在词汇理解和词汇产生过程中都会存在 AoA 效应，但由于具体任务中所涉及的加工过程有差异，其 AoA 效应的大小是不同的。根据多层次理论的观点，图画命名任务所包括从感知觉到发音输出中的各个加工阶段，AoA 效应大于词汇判断、语义范畴判断和词汇朗读任务，前两类任务涉及到额外的语义加工，因此应比词汇朗读任务产生更大的 AoA 效应(Moore et al., 2004; Catling & Johnston, 2009)。根据表征理论的观点，涉及到语义加工过程的任务均会呈现出 AoA 效应，因此图画命名、词汇判断任务和语义范畴判断均会出现 AoA 效应，而词汇朗读任务中语义加工程度较低，AoA 效应较小，其大小应与词频效应相近(Brysbaert & Ghyselinck, 2006; Kuperman, 2013)。根据映射理论的观点，在所有任务中，图画命名中对应的语义表征与语音表征间的映射程度一致性最低，AoA 效应最大；与词汇朗读任务相比，语义范畴判断与词汇判断任务中词汇的正字法-语义之间的映射关系更不一致，AoA 效应较大；与语义范畴判断相比，词汇判断任务中假词的存在导致正字法和语义之间的映射关系被突出强调，因此 AoA 效应较大(Lambon Ralph & Ehsan, 2006; Catling & Elsherif, 2020)。

无论是词汇理解还是词汇产生过程，均有可能涉及到语义加工和语音信息的加工。基于表征理论的观点，对不同任务类型中是否涉及语义或语音信息进行分类，对比是否存在 AoA 效应的不同。研究发现图画命名任务中存在显著的 AoA 效应，但词汇朗读任务却不存在 AoA 效应，表明任务的语义相关度影响了 AoA 效应的大小(Lambon Ralph & Ehsan, 2006)。与此相比，也有研究在语义范畴判断任务中发现其 AoA 效应小于图画命名任务，可能是由于语音加工卷入程度不同导致的(Catling & Johnston, 2006c)。与语义或语音相关度较低的任务相比，具有高语义或语音相关度的任务 AoA 效应更大。

1.2.2 文字系统

通过比较已有研究发现，AoA 效应具有跨语言的一致性，研究者在英语(Carroll & White, 1973a)、西班牙语(Izura & Hernández-Muñoz, 2017)、德语(Kauschke & Von Frankenberg, 2008)、荷兰语(Ghyselinck, Custers, Brysbaert, 2004)、汉语(白学军 等, 2010; 陈宝国 等, 2007; Zhang et al., 2022)等各类语言中均发现了 AoA 效应。根据文字系统可以将发现 AoA 效应的语言分为两类：多数印欧语系使用的是表音文字系统，例如英语、德语、法语、西班牙语和荷兰语等，而汉语使用的则是表意文字系统。表音文字的字符直接与声音相关，单词根据拼读规则标记发音，具有较规律的正字法-语音映射关系(Grainger & Holcomb, 2009)，但与语义无直接关联。而表意文字的字符通常与意义相关，汉字以形声字为主，其中形旁反映了汉字的语

义类属，同时声旁也有助于语义提取(张积家 等, 2014)，使得汉语中正字法-语义映射更为规则(Chang & Lee, 2020)。虽然汉语的声旁能在一定程度上传递语音，但与表音文字相比，汉语声旁的读音并不固定(Chen et al., 2007)，表意文字仍以表达意义为主要特点。因此，表音文字系统中正字法和发音之间的映射关系较为直接，正字法和语义、语音和语义之间的映射关系可能存在更高的任意性，而表意文字系统中的正字法和语义之间的映射关系较为直接，正字法和语音、语音和语义之间的映射关系可能存在更高的任意性(Weekes, 2011; Yum & Law, 2019; 白学军 等, 2007)。研究者通过比较日语中存在不同映射关系的两种文字系统(属于表意文字的汉字与属于表音文字的假名)，发现与假名相比，使用汉字完成词汇朗读任务具有更大的 AoA 效应，表明文字系统调节了 AoA 效应大小，支持了任意映射理论的观点(Havelka & Tomita, 2006)。Zevin 和 Seidenberg(2002)也认为在表音文字中正字法-语音的之间的映射更加规则，不足以产生可靠的 AoA 效应。与此相比，汉语中词汇的正字法-语音之间的映射更不规则，一系列研究在词汇朗读任务中发现了可靠的 AoA 效应(陈宝国 等, 2004; 陈永香, 朱莉琪, 2015)。

1.2.3 词频

词频描述了语料范围内词汇出现的频率，是较早被关注的语言学变量之一，影响词汇的识别和处理速度(Brown & Watson, 1987; Levitt & Healy, 1985)。累积频率假说(Cumulative Frequency Hypothesis)认为 AoA 效应主要来源于词汇的使用频率高低，即早期获得的词汇由于使用频率高，在语言任务中表现出了更短的反应时间(Gerhand & Barry, 1999a)。研究表明，在图画命名(Lambon Ralph & Ehsan, 2006)、词汇朗读(Gerhand & Barry, 1998)、词汇判断(Wilson et al., 2013)等任务中观察到了词频与 AoA 间存在交互作用，表现为高频词下 AoA 效应更小，而低频词下 AoA 效应更大。Ellis 和 Lambon Ralph(2000)从联结主义的观点出发，利用计算建模方法模拟了词汇习得过程，通过调整输入输出模式对的顺序模拟 AoA 效应，按照事先设定的频率反复训练已经输入的模式，考察词频在 AoA 效应中的影响，结果表明输出的准确性受输入早晚的影响，且这一影响在高频和低频词中存在差异。表征理论认为 AoA 效应部分来源于词频效应，在语义网络的建构过程中，系统状态与训练次数同时影响网络学习的结果(Brysbaert & Ghyselinck, 2006)。

综上，目前针对 AoA 效应的研究绝大多数为行为研究，仅有少量研究采用脑电技术考察了 AoA 效应发生的时间进程(Bakhtiar et al., 2016; 娄昊 等, 2019)或采用功能性磁共振成像方法考察了 AoA 效应的神经机制(Weekes et al., 2004)。针对行为研究的发现，在理论层面，

研究者从不同的视角关注了 AoA 效应发生的认知机制，早期理论主要从表征或不同表征之间的映射关系进行解释，最近的整合理论则综合考虑了这两个方面。已有研究发现 AoA 效应存在两种不同的来源，主要依据早晚获得词汇在各类任务之间的差异是否达到了显著水平，总结了影响 AoA 效应的因素(见综述 Dirix & Duyck, 2017; Elsherif & Catling, 2022)。传统分析未厘清各种影响因素对 AoA 效应的贡献大小，未对比各种影响因素对 AoA 效应的影响，因此，本研究主要针对行为研究指标，采用三水平元分析深入系统地考察 AoA 效应的认知机制。

传统元分析是两水平结构数据的汇总结果(Pastor & Lazowski, 2018)，其结果变异来源于两部分，包括个体数据的变异和研究之间的异质性。元分析需要满足效应量独立的基本假设，如果同一个研究包括多个效应量，这些效应量之间会因样本、实验环境等因素而存在相关性，研究者通常采用其中一个效应量或者取多个效应的均值，忽略效应量之间的差异(Lipsey & Wilson, 2001)。基于对 AoA 效应行为研究的梳理，我们发现多篇文章都涉及多项目实验任务与多个效应量(Catling & Johnston, 2006b; Holmes & Ellis, 2006; Raman, 2018)，各个效应量之间并不独立。为了解决同一研究内多个效应量之间存在相关性的问题，且最大限度地利用已有研究中的效应量(Cheung, 2014)，我们采用了三水平元分析，在传统元分析的基础上增加了一个变异水平，即来自同一样本多个效应量之间的变异。

基于各理论的争论焦点，我们在元分析中系统地分析和对比目标认知过程、具体的任务类型、任务与语义或语音信息的相关度等因素对 AoA 效应的调节作用。其次，根据任意映射理论的观点，不同文字系统各层表征间的映射关系不同，因此我们会跨研究对比不同文字系统(表音文字和表意文字)中的 AoA 效应量。此外，基于词汇的语言学特点，研究者发现 AoA 效应会受到词汇词频的影响，因此同时关注词频对 AoA 效应的调节作用。

2 研究方法

2.1 文献检索及纳入标准

中文文献来源于中国知网(包括期刊全文数据库、硕博论文全文数据库和会议论文数据库等)、万方数据知识服务平台及维普中文期刊服务平台, 检索主题词或关键词为“词汇获得”或“词汇获得年龄”或“词汇获得早晚”或“词汇获得年龄效应”或“词汇习得年龄”。英文文献从 Web of Science(包括 SCI、SSCI、A&HCI、ESCI、MEDLINE、Preprint、ProQuest 等)、 Science Direct、Scopus 数据库中获得，检索关键词为“age of acquisition”或“AoA”或“age of acquisition effect”或“AoA effect”。所有文献检索的时间范围是 2024 年 5 月 15 日以前。文献筛选流程图

见图 1。

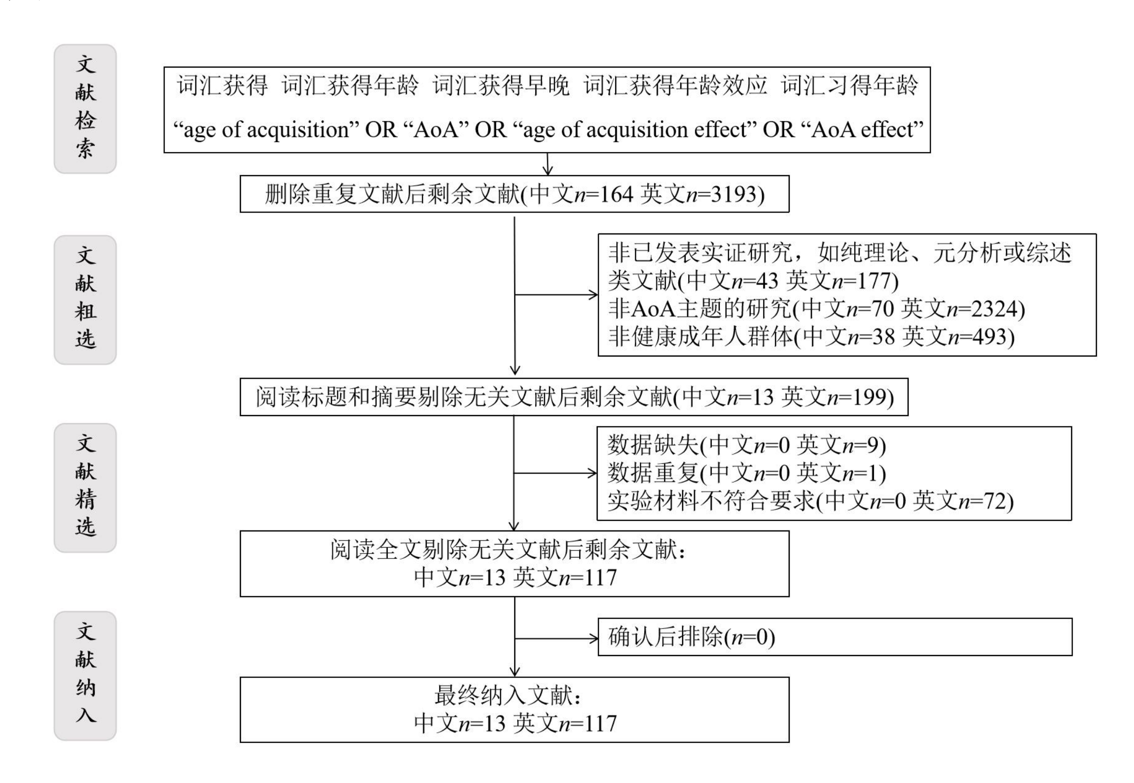


图 1 元分析文献筛选流程

对于检索到的相关研究，主要按照以下标准进行筛选，考虑是否纳入元分析：(1)需为词汇获得年龄效应主题的已发表实证研究，排除会议论文、学位论文、纯理论和文献综述类文章；(2)排除无关主题文章，如二语习得研究、未涉及词汇加工任务的词汇库或图画库构建研究；(3)研究对象为健康成年人群体，排除儿童、老年人、语言障碍患者、阿尔兹海默症患者等特殊被试群体；(4)实验材料为名词、动词或形容词，排除实验材料为人名、情绪词的研究；(5)数据重复发表的仅取其一；(6)排除没有报告反应时的研究，如单词默读任务、延迟命名任务、知觉辨认任务等。所选文献中需详细报告不同词汇获得年龄条件下反应时的 Hedge's g 值或能够转化为 Hedge's g 值的反应时均值和标准差、F 值、t 值、β值和 R^2。最终得到符合上述标准的文献 130 篇，其中中文文献 13 篇，英文文献 117 篇，共 271 项效应量。

2.2 文献评估与编码

为了对原始文献的结果进行定量综合，本研究参照 Lipsey 和 Wilson(2001)的方法对文献进行编码，并利用 SPSS26.0 软件计算了 Kappa 一致性系数。主要编码的信息和一致性检验结果包括：(1)文献信息(作者名、发表时间); (2)样本量(Kappa = 0.89); (3)文字系统(表音文字 vs.表意文字)(Kappa = 0.84); (4)目标认知过程(感知和产生)(Kappa = 0.89); (5)实验任务(图画命名、词汇判断、词汇朗读、语义范畴判断和其他任务)(Kappa = 0.82)(6)任务的语义相关度

(高 vs.低)(Kappa = 0.82)；(7)任务的语音相关度(高 vs.低)(Kappa = 0.82); (8)材料是否控制词频(控制、未控制和独立)(Kappa = 0.88)；(9)效应量(反应时均值和标准差、F 值、t 值、β值、R^2)(Kappa = 0.63)。

编码遵循如下原则：(1)每个独立样本的一个主效应编码一次，若一篇文献包含多个独立样本则分别编码。(2)若同一样本报告了多个显著的效应量，则每个效应量单独编码。(3)若被试包括不同组(如青年组/老年组，健康成年人组/阅读障碍者组)，则仅对青年组和健康成年人组的样本量以及相关信息进行编码。(4)表音文字系统包括英语、法语、西班牙语、德语、荷兰语、波兰语、印地语、冰岛语、土耳其语，以及韩语和一种印度南部的达罗毗荼语系语言坎纳达语(Kannada)。研究者认为韩语的拼写规则与其发音紧密相关(Han & Kim, 2017; Han et al., 2021)，而坎纳达语属于拼写一致的语言(Nishimoto et al., 2005)，故而统一为归属于表音文字系统。表意文字系统主要为汉语研究。由于日语文字系统的复杂性，在文字系统变量的调节效应检验中排除了日语材料相关研究。(5)任务的语义相关度依据实验任务类型以及实验设计的操纵变量进行编码。如引言所述，相比图画命名任务、词汇判断任务和语义范畴判断任务，词汇朗读任务的语义加工深度较低(但并不意味着词汇朗读任务不涉及语义加工)，故而一般词汇朗读任务的语义相关度编码为低，图画命名任务、词汇判断任务和语义范畴判断任务的语义相关度编码为高。部分词汇判断任务的实验材料采用了不符合正字法规则的假词，这些任务的效应量参与了主效应检验和文字系统、目标认知过程的调节效应检验，不参与任务类型、任务语义相关度和任务语音相关度的调节效应检验。部分词汇朗读任务要求被试在朗读的基础上对词汇进行语义加工，如仅对真词进行朗读，而不对假词进行反应；呈现动词原型要求被试说出其现在进行时或过去式(Morrison et al., 2003)，这些任务不进行任务类型、任务语义相关度和任务语音相关度的调节效应检验。(6)任务的语音相关度同样依据实验任务类型以及实验设计的操纵变量进行编码，其中涉及发音的图画命名和词汇朗读任务的语音相关度为高，词汇判断任务和语义范畴判断任务的语音相关度为低。

文献编码工作由两位心理学专业研究生独立完成，在完成后进行交叉检验。对于两位编码者存在分歧的部分，由两位编码者与通讯作者讨论后确定。最后，纳入元分析的 130 篇文献共报告了 271 个效应量，10697 名被试(见附录 1)，文献发表于 2024 年 5 月 15 日以前。在同一项实验中，效应量数最少的是 1 个，最多的是 4 个。

文献质量评估采用美国国立卫生研究院(National Institutes of Health, NIH)提供的纵向和横断研究质量评估工具(Quality Assessment Tool for Observational Cohort and Cross-Sectional

Studies)，排除了 4 个明显不适用于 AoA 效应研究的纵向研究质量评估题项后，本研究按照符合记 1 分、不符合标准记 0 分的标准对所纳入的各项研究进行了质量评分(National Institutes of Health, 2018)，总分介于 0~10 之间。研究质量评分结果见附录 1，评分越高表明文献质量越好。

2.3 元分析过程

本研究采用 R 软件和 metafor 数据包进行数据处理与分析。元分析过程包括：在获得每个研究效应量的基础上，首先检验文献是否存在发表偏倚；其次采用三水平元分析模型进行异质性检验(Heterogeneity test)、主效应检验和调节效应检验，调节变量的分析采用亚组分析检验分类变量的结果是否显著。

2.3.1 效应量计算

本研究使用 Hedge's *g* 值作为效应量，以 0.10、0.30 和 0.50 为临界值，分别判定小、中和大的效应量(Cohen, 1992)。多元回归分析研究通过标准回归系数(β)、R^2计算效应量大小，因素设计的研究多数通过反应时均值、标准差、样本量计算效应量大小，少量研究通过t值、*F* 值转换为 Hedge's *g*(Ben-Shachar et al., 2020)。采用 R 软件和 metafor 数据包进行三水平元分析(Assink & Wibbelink, 2016)。

2.3.2 元分析模型选择

元分析纳入的 130 篇文献中有 70 篇报告了两个或以上的效应量，并且其中有 141 个效应量与 1 个或多个其他效应量来自同一组样本，这些效应量彼此存在依赖性(Cheung, 2014)，违背了传统元分析方法认为效应量相互独立的假设。若采用传统元分析方法可能导致元分析结果被高估(Lipsey & Wilson, 2001)，而三水平元分析可以解决这一问题。三水平元分析将方差来源分解为三个水平：抽样误差引起的效应量变异(水平 1：抽样方差)、从同一研究提取的效应量之间的变异(水平 2：研究内方差)、从不同研究所提取的效应量之间的变异(水平 3：研究间方差)(Cheung, 2014)，可以提取原始研究的全部效应量，从而最大程度保留信息，提高统计检验力。基于上述原因，本研究将使用三水平元分析模型(郭英 等, 2023; 孟现鑫 等, 2023; 孟现鑫 等, 2024; 赵子卿 等, 2024; 诸彦含 等, 2024)进行异质性检验、主效应检验、发表偏倚检验以及调节效应检验。

2.3.3 异质性检验与调节效应检验

本研究通过 Q 检验评估总体的异质性，并对研究内方差(水平 2)和研究间方差(水平 3)进行单侧对数似然比检验(one tailed log likelihood ratio tests)进一步确定异质性分布。当检验

显著时，说明存在异质性，以 25%、50%、75%的 I^2 值作为区分异质性低、中、高的界限(Higgins et al., 2003)，并进行调节效应检验以进一步确定异质性的来源(孟现鑫 等, 2024; Assink & Wibbelink, 2016; 赵子卿 等, 2024)。本研究将调节变量作为协变量加入三水平元分析模型(孟现鑫 等, 2023)，以估计其调节效应大小。研究中的调节变量均为分类变量，包括文字系统、认知过程、任务类型以及词频。为了保证调节效应检验结果的代表性，分类调节变量每个水平的效应量个数应不少于 5(Card, 2016)，因此本研究在任务类型、任务的语义相关度和任务的语音相关度的调节效应检验中，仅分析图画命名、词汇判断、词汇朗读和语义范畴判断任务这四类。对于其他任务类型，如渐进式去掩蔽任务[1](Progressive Demasking, Gilhooly & Logie, 1981b)、图画验证任务(Catling & Elsherif, 2020)、物体真实性判断任务(Catling & Johnston, 2009)、视觉持续时间阈限测定任务[2](Visual Duration Threshold Task, Dent et al., 2007)、模糊字辨认任务(陈宝国 等, 2010)、词汇回忆任务(Gilhooly & Gilhooly, 1979)等的效应量仅进行了主效应检验、文字系统和目标认知过程的调节效应检验。

2.3.4 发表偏倚和敏感性检验

如果纳入元分析的文献不能全面代表该领域研究的总体结果,则会出现发表偏倚(Publication bias)(Rothstein et al., 2005)。为了检验是否存在发表偏倚，本研究运用漏斗图(Funnel plot)、Egger’s 回归法、失安全系数 N 进行检验。漏斗图是由各个效应量转化而成的可视化散点图，若不存在发表偏倚，数据应左右对称分布、集中在中上部，汇集成一个大致对称的漏斗形状(Light & Pillemer, 1984)。Egger’s 回归检验中 p 值不显著，则不存在严重的发表偏倚(Egger et al., 1997)。失安全系数 N 大于 $5 \times k$+10(k 为效应量数量)，则可以忽略发表偏倚(Rothstein et al., 2005)。若存在发表偏倚，则采用剪补法(Trim and fill method)进行偏倚矫正(Duval & Tweedie, 2000)。若剪补后的效应量未发生显著变化，则可认为该元分析结果受发表偏倚影响较小(Duval & Tweedie, 2000)。

为评估本元分析结果的稳健性，本研究采用“去一法”剔除对元分析结果可能产生显著影响的异常值，并重新进行三水平元分析以衡量异常效应量和异常研究的影响。“去一法”通过逐个剔除纳入的效应量和原始研究，并重新进行三水平元分析，直到所有的效应量和原始研究均被剔除过(Dodell-Feder & Tamir, 2018)。

---

[1] 在渐进式去掩蔽任务中，一般是在目标单词覆盖了图案掩蔽，随着时间掩蔽刺激逐渐减少，直至被试能够正确识别出单词，测量被试正确识别单词的反应时间。

[2] 在视觉持续时间阈限测定任务中，一般都是变化词汇刺激的呈现时间，从几十毫秒开始不断增加直至被试能够正确识别词汇，记录能正确识别词汇刺激的最低阈限时间。

3 研究结果

3.1 异质性检验

AoA 效应元分析Q检验的结果显示，Q值达到显著水平，$Q(270) = 1644.16$，$p < 0.001$，即本研究纳入的各个效应量之间存在异质性。从同一研究提取的效应量之间的变异(水平 2 方差) ($\sigma^2 = 0.05, p < 0.001, I^2 = 42.84\%$)和从不同研究提取的效应量之间的变异(水平 3 方差) ($\sigma^2 = 0.05, p < 0.001, I^2 = 43.17\%$)均显著。这两种变异的总$I^2$值为 86.01%，即观察变异中 86.01%由效应量的真实差异造成，13.99%由随机误差造成，表明研究间的变异存在组间误差干扰，各研究间存在较高异质性，意味着文字系统、认知过程、任务类型以及词频可能调节了词汇获得年龄效应。因此，有必要对相关调节变量进一步分析。

3. 2 发表偏倚和敏感性检验

Egger's 回归法的结果精度系数为负数($t = -4.31, df = 269, p < 0.001$)，截距为 0.64，95%CI 为[0.55, 0.73]，表明漏斗图存在显著的不对称性。漏斗图(见图 2)显示实际观测的效应量集中于 0.46 左右，标准差在 0.15 左右，呈不对称分布。为使漏斗图呈对称分布，我们采用了剪补法，结果显示在漏斗图左侧需要剪补 0(SE = 9.41)个效应量，表明没有明显遗漏小样本或不显著的研究。失安全系数 $N = 32491$，远大于 $5k + 10$ 这一临界值，说明当前的效应结果十分稳健。因而当前主效应结果受发表偏倚的影响较小。

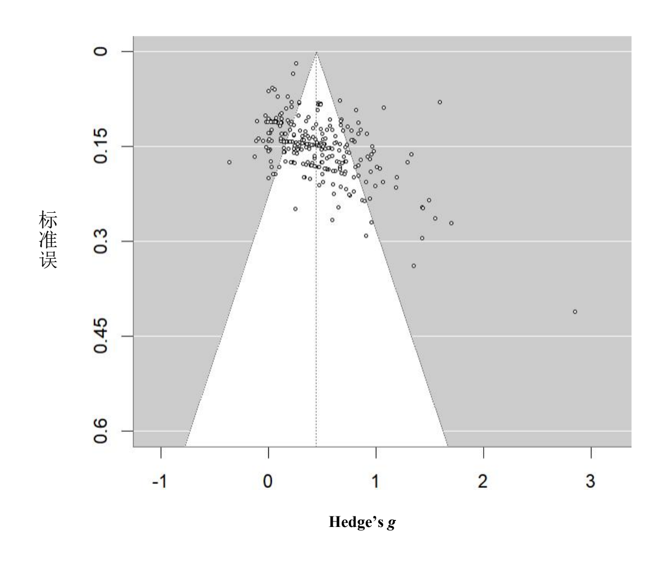


图 2　AoA 效应量分布漏斗图

(注：黑点代表实际观测的效应量, 中间垂直的虚线为主效应量; 两边虚线为 95%的置信区间。)

采用去一法逐个剔除纳入的效应量并重新进行三水平元分析，结果显示，剔除了 Karimi

和 Diaz(2020)报告的一个效应量后，AoA 效应的效应量最低(Hedge's $g = 0.46$, $df = 269$, $p < 0.001$)；剔除了 Morrison 和 Gibbons(2006)实验 2 报告的一个效应量后，AoA 效应的效应量最高(Hedge's $g = 0.47$, $df = 269$, $p < 0.001$)。敏感性分析结果发现，剔除前与剔除后重新计算的主效应在显著性上一致，且都属于中等大小的效应量，故而当前元分析结果受异常值影响小。

3.1.3 调节效应检验

为检验不同研究中效应大小的可能来源，使用亚组分析的方式进一步考察可能调节 AoA 效应的各种分类变量，其中各亚组至少含有 21 个效应量，高于蓝媛美等(2022)建议的各亚组至少含有 3 个效应量。

利用亚组分析检验各调节变量对 AoA 效应是否存在显著影响(结果如表 1 所示)。目标认知过程的调节效应不显著，$F(1, 269) = 0.08$, $p = 0.781$。任务类型的调节效应显著，$F(3, 224) = 9.32$, $p < 0.001$，其中图画命名任务(Hedge's $g = 0.55$)与词汇判断的效应量(Hedge's $g = 0.59$)均显著大于词汇朗读任务(Hedge's $g = 0.34$)和语义范畴判断任务(Hedge's $g = 0.34$)，而图画命名任务与词汇判断任务的效应量没有显著差异($t(131) = –0.57$, $p = 0.569$)。任务的语义相关度的调节效应显著，$F(1, 219) = 14.62$, $p < 0.001$，其中语义相关度高时的 AoA 效应量(Hedge's $g = 0.53$)显著大于语义相关度低时的效应量(Hedge's $g = 0.34$)。任务的语音相关度对 AoA 效应的调节效应不显著，$F(1, 219) = 1.27$, $p = 0.261$。文字系统的调节效应不显著，$F(1, 265) = 2.80$, $p = 0.096$，表意文字的效应量数值上小于表音文字的效应量($\beta = –0.11$(SE = 0.07), $t = –1.67$, $p = 0.096$, 95%CI = [–0.24, 0.02])。实验材料是否控制词频的调节效应显著，其中以未控制为参照水平，控制了词频(Hedge's $g = 0.53$)和将词频作为独立变量(Hedge's $g = 0.58$)的效应量均显著高于未控制词频条件(Hedge's $g = 0.37$)，而控制了词频和将词频作为独立变量的研究效应量之间没有显著差异($t(166) = –0.67$, $p = 0.503$)。

对于将词频作为独立变量的 45 项研究，其中有 6 项研究发现了显著的 AoA 和词频的交互作用，并且任务类型主要集中于词汇判断任务，另有图画命名任务和词汇朗读任务各一项。在报道了高频词和低频词条件下 AoA 效应(均为交互作用显著)的研究中，发现对比低频词条件，高频词条件下的 AoA 效应显著减小($\beta = 0.80$(SE = 0.33), $t(7) = –2.41$, $p = 0.042$, 95%CI = [–1.56, –0.04])。

表 1 词汇获得年龄效应的调节效应检验

调节变量	k	Intercept/ Hedge's g(95%CI)	β(95%CI)	F 值	p	研究内方差	研究间方差
目标认知过程				0.08	0.781	0.05	0.05
感知	105	0.46(0.38, 0.53)***					
产生	166	0.47(0.41, 0.54)***	0.01(–0.08, 0.10)				
文字系统				2.80	0.096	0.05	0.05
表音文字	220	0.50(0.45, 0.56)***					
表意文字	47	0.47(0.41, 0.54)***	–0.11(－0.24, 0.02)				
任务类型				9.32	< 0.001	0.03	0.06
词汇朗读	74	0.34(0.25, 0.43)***					
语义范畴判断	21	0.34(0.17,0.51)***	–0.03(–0.19, 0.14)				
图画命名	79	0.55(0.47, 0.63)***	0.22(0.10, 0.33)***				
词汇判断	54	0.59(0.49, 0.69)***	0.25(0.13, 0.37)***				
语义相关度				14.62	< 0.001	0.04	0.05
低	70	0.34(0.25, 0.43)***					
高	151	0.53(0.47, 0.60)***	0.20(0.10,0.30)**				
语音相关度				1.27	0.261	0.05	0.04
低	72	0.52(0.44, 0.61)***					
高	149	0.47(0.41, 0.54)***	–0.06(–0.16, 0.04)				
词频				6.93	0.001	0.04	0.04
未控制	103	0.37(0.29, 0.44)***					
控制	123	0.53(0.45, 0.61)***	0.17(0.07, 0.28)**				
独立	45	0.58(0.42, 0.75)***	0.22(0.07, 0.37)**				

注：$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$。

4 讨论

本研究采用三水平元分析的手段，考察了 AoA 效应的总效应大小以及可能的调节变量的影响。元分析表明 AoA 效应量较大，早获得词比晚获得词的加工速度快、正确率高，表明在各实验任务和各语系中，词汇获得年龄都是一个稳定的影响词汇加工的重要因素。元分析结果表明任务类型、词频、任务的语义相关度等三个重要变量在 AoA 效应中起调节作用，而目标过程、文字系统和任务的语音相关度对 AoA 效应量无显著的调节作用，我们将结合三类理论观点以及整合理论阐述上述各因素在 AoA 效应的作用。

4.1 任务类型和语义相关度在 AoA 效应中的调节作用

元分析结果表明目标加工为词汇理解还是词汇产生过程，对于 AoA 的效应量无显著影响，而任务类型显著地影响了 AoA 效应量的大小。具体而言，图画命名、词汇判断、语义范畴判断和词汇朗读四种任务中 AoA 效应的效应量均达到了显著水平($p < 0.001$)。进一步分析表明，图画命名与词汇判断任务的 AoA 效应量显著大于词汇朗读和语义范畴判断任务，而图画命名与词汇判断之间($t(131) = 0.57, p = 0.569$)以及语义范畴判断与词汇朗读任务之间的 AoA 效应量差异不显著($t(93) = -0.32, p = 0.740$)。词汇朗读任务主要反映了语音加工过程，语义范畴判断任务则主要涉及语义加工过程(Balota & Chumbley, 1984)。该任务类型的效应量结果显示，AoA 同样参与了语义加工阶段，表明 AoA 可能同时影响了语义加工阶段和语音加工阶段。此外，其他涉及字形或者知觉加工层面的任务(如渐进式去掩蔽任务、模糊字辨认任务、视觉持续时间阈限任务等)也报告了显著的 AoA 效应，说明 AoA 同样影响了知觉加工阶段。这一结果支持了多层次理论，即 AoA 对词汇加工的影响涵盖了知觉、语义和语音加工等各个阶段。

任务类型的调节效应分析表明这一变量对 AoA 的效应量的影响不显著，这一结果模式与具体 AoA 效应大小的结果对比不一致。在具体的 AoA 效应大小上，Elsherif 等(2023)分别计算了不同任务类型中 AoA 效应大小的平均值，发现图画命名任务中的效应值(107ms)显著大于词汇判断任务(45ms)和词汇朗读任务(23ms)，且词汇判断和词汇朗读之间的差异也达到了显著水平。这一综述中所采取的指标为 AoA 效应的具体大小，在跨任务之间比较时，AoA 效应的具体大小会受到实验材料和被试样本量等因素的不同影响。而本研究中采用的指标是可以进行跨任务、跨研究、跨群体比较的效果量 Hedge's *g* 值，这一指标排除了样本量大小的干扰，能够较为直观地比较不同研究的结果。元分析中在文献纳入过程中排除了学术论文等可能存在发表质量的文章，同时仅保留了以名词、动词为研究对象的研究，使得研究间更

具有可比性。因此，我们认为元分析的指标更为可靠地反映了 AoA 效应在各个任务类型之间的差异模式。

研究结果表明图画命名任务和词汇判断任务的 AoA 效应量相当，均大于语义范畴判断任务和词汇朗读任务中的 AoA 效应量。一种可能的解释是从表征理论观点出发：无论是反映词汇理解的词汇判断任务还是反映词汇产生的图画命名任务，这两种任务中都涉及到语义系统的加工，因而产生了相当的 AoA 效应量，而词汇朗读任务中不涉及到语义系统的加工，仅涉及到音韵系统的加工(Cortese et al., 2018; Elsherif et al., 2020; Kuperman, 2013)。与此解释一致，在元分析中我们观察到了语义相关度的显著影响，语音相关度则没有产生显著影响。在效应量上，图画命名任务和词汇判断任务呈现出一个中等水平的效应量，而词汇朗读任务呈现出一个较小的效应量，表明 AoA 效应不仅受到任务语义相关度的影响，而且也受到任务语音相关度的影响(Elsherif et al., 2023)。

第二种可能的解释来自于多层次理论的观点。词汇判断任务和图画命名任务中都涉及到语义、音韵和正字法系统的参与，而语义范畴判断任务主要涉及正字法和语义系统的加工，词汇朗读任务中涉及到正字法和音韵系统的参与。从加工过程上来说，词汇判断和图画命名两类任务所包括的加工阶段比语义范畴判断任务和词汇朗读任务多，因而产生了更大的AoA效应量，后两种任务的AoA效应量不存在显著差异。尽管如此，与词汇判断任务相比，研究者认为图画命名任务中同时包括了概念水平的加工(Catling & Johnston, 2009)，那么应该在图画命名任务中观察到一个更大的AoA效应。元分析结果与此不一致，提示可能有其他因素同时影响了AoA效应的效果量大小。

我们对AoA效应的脑电研究做了元分析[3]，结果显示AoA可以稳定调节N400波幅大小(Hedge’s $g = 0.45$, $p < 0.001$, 95%CI = [0.29, 0.62])，早获得词诱发的N400波幅小于晚获得词，表明对晚获得词的加工更慢，其语义提取也更难，说明AoA效应的产生需要语义加工过程的参与，支持了表征理论。尽管如此，N400波形是否仅仅反映了语义加工，与语音加工无关？研究者需要进一步深入考察脑电和行为指标之间的关系，探索大脑与行为之间的关系。

根据任意映射理论的观点，AoA效应来自于词汇获得与发展过程中词汇网络可塑性的变化，各类表征系统之间的映射关系会影响AoA效应，规则的映射关系能在一定程度上改善网络可塑性降低所带来的加工劣势。尽管词汇判断、语义范畴判断、图画命名和词汇朗读分别

[3] 脑电的元分析中纳入了 11 篇文献，共 17 个效应量，421 名被试。由于 ERP 文献数量较少，同时研究结果发现的波幅成分较为分散，包括 N400(Bakhtiar et al., 2016; Cuetos et al., 2009; Räling, 2015; Yum & Law, 2018)、LPC(Bakhtiar et al., 2016; 白学军 等, 2010)、SN(Adorni et al., 2013)、P2(Fargier & Langanaro, 2020)、P3(Tainturier et al., 2005)、N540(白学军 等, 2010)，以及仅给出脑电波幅差异显著时窗并未说明其具体成分的研究(Laganaro & Perret, 2011; Laganaro, 2014; Perret et al., 2014; Valente et al., 2014)。基于元分析的标准，我们认为较可靠的结果仅反映在 N400 上。相关结果见附录 2。

对应于词汇理解和词汇产生过程，但理解和产生过程中所用的正字法、语义和语音表征系统是相同的。词汇朗读任务依靠词汇层面正字法-语音之间的映射关系完成，无论是表音文字还是表意文字，都存在一定的正字法规则使得个体可以通过正字法信息获取其语音信息，正字法-语音的映射存在一定的规则(Monaghan & Ellis, 2002a; Raman, 2006)。词汇判断和语义范畴判断任务涉及到正字法-语义表征之间的映射关系，正字法与语义的联系并不直接，难以通过正字法信息与语义建立直接的联系，因此正字法-语义间的映射一致性较低(Chen et al., 2007)。图画命名任务涉及到语义-语音表征，语义与语音之间的映射很大程度上是任意的，词汇发音与其所代表的意义并没有必然的联系，语义-语音间映射的一致性也较低。因此，由于映射关系的规则性不同，图画命名任务和词汇判断任务中的效应量相似，而语义范畴分类任务和词汇朗读任务中的效应量小于前两类任务。与多层次理论和表征理论相比，任意映射理论能一致地解释本研究所发现的各种任务对比中AoA效应量的模式。

也有研究者持不同的观点，Elsherif等人(2023)认为与词汇判断任务相比，图画命名任务中语音和语义之间映射关系的规则性和系统性更低，根据任意映射理论的观点，AoA效应在词汇判断任务中应该更小。当把任务在语义相关度和语音相关度两个维度上进行区分后进行分析，结果表明语义相关度的调节作用显著，而语音相关度的调节作用不显著。这一结果表明，与不涉及语义的任务相比，涉及到语义加工的任务其AoA效应量更大，但是任务中是否涉及语音加工并不影响AoA效应的大小，表明语义而非语音在AoA效应中起了重要的调节作用。已有研究认为AoA效应可能发生在多个加工阶段，包括概念和语义加工以及语音加工(Brysbaert & Ellis, 2016; Brysbaert et al., 2017; Elsherif et al., 2023; Hernandez & Li, 2007; Johnston & Barry, 2006; Juhasz, 2005)，然而语音相关度的元分析结果表明语音能够起到的作用是有限的。尽管在词汇朗读任务中发现了显著的AoA效应，表明AoA会影响语音加工阶段。但AoA效应量并不受到语音加工深度的影响。这一研究发现提示研究者，考察AoA效应的认知机制可能需要从语义表征或语义网络的角度出发，探索语义层面的影响以及语义层与其他水平之间的联系。

4.2 词频对AoA效应的调节作用

元分析结果显示：与未控制词频相比，研究中匹配了早晚获得词的词频或者将词频作为一个变量进行独立操纵时，其 AoA 效应更大。同时，元分析在考察多元回归和因素设计两类研究中的 AoA 效应量时发现，多元回归分析研究得到的 AoA 效应量显著小于因素设计研究得到的效应量大小($t(269) = 4.47, p < 0.001$)。未控制词频的研究均为多元回归分析，而匹配了早晚获得词的词频或将词频作为一个变量进行独立操纵的研究采用了因素设计。我们认

为两类研究中所表现出的 AoA 效应差异来源于实验材料的不同。在多元回归分析中，研究者同时考虑词汇的 AoA 和词频的影响，这两个指标之间存在高相关(Morrison et al., 1997；Zevin & Seidenberg, 2002)。回归分析中当多个变量之间存在高度相关时，可能会使得某些预测变量的作用难以体现，从而导致 AoA 效应在该方法中未能完全显现(Morris, 1981)。陈宝国等(2004)在汉语研究中采用了回归分析方法，考察了 AoA 与词频等变量对汉语双字词命名的影响，结果表明词频影响了词汇的命名反应时，未发现 AoA 对词汇朗读产生影响。在因素设计中，词频和 AoA 被当作独立的变量进行严格操纵，其余可能影响词汇加工的变量需进行匹配。为了独立地操纵这两类指标，所选择实验材料的词频和 AoA 指标的变化范围均局限在一定范围内，不能代表词汇的整体分布(Brysbaert & Ellis, 2016)。相比而言，在回归分析中，词频和 AoA 指标作为连续变量，所使用的材料更接近词汇实际整体。

研究者对于 AoA 效应和词频效应的认知机制存在争论，累积频率假说认为 AoA 效应源自早习得词汇在整个生命周期中有更高的累积频率(Balota et al., 2004; Zevin & Seidenberg, 2004)，表征理论则认为 AoA 与词频效应共享认知机制(Brysbaert & Ellis, 2016; Brysbaert & Ghyselinck, 2006)，也有研究结果表明 AoA 和词频独立地影响词汇加工过程(Cortese & Khanna, 2007; Gerhand & Barry, 1998; 陈宝国 等, 2004)。元分析发现在词频作为一个变量进行独立操纵的因素设计研究中，23.8%的研究报告了 AoA 与词频的交互作用是显著的，主要表现为低频词比高词频词汇表现出更大的 AoA 效应，表明词频与 AoA 可能共同地影响了词汇加工中的某一个加工阶段，比如词汇的语义提取，或者两个变量分别影响了两个加工阶段，但是这两个加工阶段之间存在密切联系。然而也有研究未发现 AoA 与词频的交互作用显著(Burani et al., 2007; Catling & Elsherif, 2020; Dewhurst & Barry, 2006; Fiebach et al., 2003; Gerhand & Barry, 1998; Ghyselinck, Lewis, Brysbaert, 2004; Meschyan & Hernandez, 2002; Wilson et al., 2013; 陈宝国 等, 2004; 王丽红 等, 2010)，支持 AoA 效应独立于词频存在。

AoA 效应和词频效应可能是独立的(Cortese & Khanna, 2007)或部分独立的，两者具有不同的认知机制。在图词干扰范式中，研究者操纵了干扰词与目标词的语义和语音相关度，发现 AoA 仅与语义相关度存在交互作用，而不与语音相关度发生交互，但是词频与语义相关度和语音相关度均存在交互作用，表明 AoA 和词频共同影响了口语词汇产生中的词汇选择过程，但仅有词频影响了音韵编码过程(Zhang et al., 2022)。利用高时间分辨率的脑电技术，研究者要求被试判断视觉呈现的词汇中是否包含特定的字母，结果发现在 180~240ms 时间窗内左枕颞区的早期选择性负波(Selection Negativity, SN)的波幅表现出 AoA 和词频的交互作用，表明 AoA 和词频共同影响了词汇加工的早期过程，可能与词汇理解过程中的正字法

信息加工有关；同时在 190~220ms 时间窗内也发现在左前区域出现了负波成分(Left Anterior Negativity component, LAN)，但是词频和 AoA 独立地影响了这一波形，并未观察到显著的交互作用，表现为早习得的词汇和高频词的波幅更高。研究者认为上述模式反映了在电生理层面上，两类效应共享部分类似的脑机制(Adorni et al., 2013)。采用计算模拟技术，研究者发现在模拟中操纵词汇出现的比率，当高低词频的频率比例为 10:1 时，AoA 和词频之间存在交互作用，而当高低词频比例为 3:1 时，两个变量之间无交互作用(Ellis & Lambon Ralph, 2000)。基于以上计算模拟和 ERP 研究结果，我们认为 AoA 与词频之间是否独立或存在交互作用依赖于词频的变化范围和早晚获得词汇的词频相对指标。

4.3 AoA 效应不受文字系统的影响

我们在元分析中未发现文字系统对 AoA 效应量的影响(p = 0.102)。根据映射理论的观点，词汇加工中正字法、语义、语音等系统之间映射的规则性会影响 AoA 效应，映射关系越规则，AoA 效应越小。表意文字中正字法-语义之间的映射更加规律，而表音文字正字法-语音的映射则更加规律。考虑到在不同的词汇加工任务中，所涉及到的映射关系是不同的，元分析中进一步比较了四类不同任务中，表音文字和表意文字的 AoA 效应量。结果发现在图画命名任务中，表音文字的 AoA 效应量边缘显著地高于表意文字(图画命名: $AoA_{表音文字}= 0.57$ vs. $AoA_{表意文字} = 0.31$, $t(77) = –1.75$, $p = 0.085$)，词汇判断任务和语义范畴判断任务中，表音文字的 AoA 效应量也出现了高于表意文字的趋势(词汇判断: $AoA_{表音文字} = 0.62$ vs. $AoA_{表意文字} = 0.46$;语义范畴判断任务: $AoA_{表音文字} = 0.38$ vs. $AoA_{表意文字} = 0.37$)，相比而言，在词汇朗读任务中，表音文字的 AoA 效应量呈现出低于表意文字的趋势($AoA_{表音文字} = 0.37$ vs. $AoA_{表意文字} = 0.41$)。尽管统计上未达到显著差异，但上述趋势与映射理论的观点基本一致。表意文字与语义的关联更强，在主要涉及到正字法-语义映射的词汇判断任务，以及主要涉及到语义-语音关系映射的图画命名任务中，表意文字对语义的加工难度更低、通达更加容易，因而其 AoA 效应更小；相比而言，在主要涉及到正字法-语音映射关系的词汇朗读任务中，表音文字的映射关系更为规则，因而其 AoA 效应更小。Bakhtiar 等(2016)考察了波斯语中元音省略的现象，元音省略降低了书写难度，但导致正字法-语音映射更加任意。根据单词元音是否省略，可以认为波斯语中存在两种映射程度不同的文字。与 Havelka 和 Tomita(2006)发现文字系统调节了 AoA 效应大小不同，Bakhtiar 等人并没有在词汇判断任务中发现 AoA 与元音是否省略在反应时上的交互作用，而在反映后期加工的脑电指标 LPC 中发现映射不一致的单词 AoA 效应更大，支持了任意映射理论。fMRI 的研究更直接的指出左侧额中回与汉语中正字法-语音的映射过程有关(Cao et al., 2009)。

虽然元分析中未发现文字系统对 AoA 效应的显著影响，但我们认为并不能否定任意映射理论的观点。对此存在以下可能的解释：一是文字系统对 AoA 效应的影响较小。从表征理论来看，文字系统可能影响个体习得词汇的类型或数量(Haman et al., 2017)，但不会影响语义网络中词汇的联结方式，因此文字系统并非是 AoA 效应的调节变量。多层次理论关注语义任务具体涉及的加工阶段，虽然不同文字系统下词汇理解与产生的加工过程存在一定的差异(Lukatela & Turvey, 1994; Rubenstein et al., 1971; 彭聃龄 等, 1985)，但本文考察的四种任务较为全面的覆盖了感知与产生的各个阶段，文字系统并没能改变语言加工涉及阶段的数量，因此多层次理论同样不认为文字系统影响 AoA 效应的大小。二是该结果可能受到文字系统研究数量的影响，本研究纳入的表意文字效应量仅有 47 个，而表音文字效应量达 220 个，不同文字系统下涉及的任务类型分布也存在差异，这可能会影响调节效应的检出(孟现鑫 等, 2024)。这提示未来还需要更多的表意文字研究来探究文字系统对 AoA 效应的影响以验证本结果的稳定性。此外，即使两种语言属于同一系统，针对两类表征之间的映射规则的一致性程度可能存在较大差异。例如，正字法-语音映射程度较高的意大利语和正字法-语音映射程度较低的英语都被认为是表音文字(Wilson et al., 2012)。为了更进一步验证映射理论的观点，我们认为更合适的方式是利用文字的特点设计实验进行研究，例如 Chen 等(2007)借用中文形声字中声旁可以预测形声字的读音这一特点，操纵了从声旁到整字读音的一致性(即正字法-语音的可预测性)，探索了不同条件下 AoA 效应的大小，研究发现支持了任意映射理论的观点。

4.4 总结、不足与展望

综上，元分析结果明确地表明语义表征在 AoA 效应中的重要作用，以及不同的映射关系在不同词汇加工任务中的影响，为表征理论和映射理论提供了可能的支持证据。三种理论从不同的角度强调了 AoA 效应源自词汇加工的表征水平或表征水平之间的联结，研究者综合了上述各种理论观点，提出了 AoA 效应的整合理论。该观点认为 AoA 效应是由早晚获得词的词汇表征和表征间的联结模式共同决定的(Catling & Elsherif, 2020)。早获得词汇的语义表征更为丰富，位于语义网络的中心，对于语义网络及其对应的神经网络结构有更大的影响；晚获得词汇与其他词汇的联系较弱，同时由于网络可塑性降低，对于神经网络的变化影响也较小，因此，语义网络更适合早获得词的输入与输出，而在加工晚获得词时需要付出更多的努力以适应其特定的输入和输出关系。整合理论既关注语义表征的作用，也强调了表征间的联结强度，两者共同决定了 AoA 效应。而元分析结果则更为强调了语义表征的作用，一定程度上反映了受 AoA 影响的语义网络对词汇加工过程的影响更大。

综合上述分析结果，我们认为本研究存在一些局限与不足，有待未来研究者进一步的探索与完善。首先，影响 AoA 效应的因素有待进一步挖掘。元分析中我们选择了任务类型、文字系统等调节变量以探究 AoA 效应的发生阶段与认知机制，对于影响词汇语义加工的典型性(Räling et al., 2015)、具体性(Song & Li, 2021)，以及影响词汇语音加工的可预测性与规则性(Chen et al., 2007)等变量并未进行考察，进一步研究可以考虑纳入更多相关的变量，对 AoA 效应进行更全面细致的探索，加深对词汇获得年龄如何影响词汇加工过程的理解。第二，元分析的结果并未明确地支持任意映射理论的观点，这与已有的部分实证研究结论并不一致(Monaghan & Ellis, 2002a; Havelka & Tomita, 2006)。已有研究未系统地区分不同任务涉及到的输入与输出间的映射关系，进一步研究可以在明确任务映射关系的基础上，操纵不同文字系统以调查映射一致性程度对 AoA 效应的影响。本文根据映射关系的侧重对文字进行分类：表意文字的正字法−语义映射相比于正字法−语音的映射具有更高的映射一致性，表音文字则相反。但文字系统间映射关系的一致性存在一定混淆，表意文字的正字法−语音映射程度并不一定低于表音文字。同一文字系统内不同文字间映射一致性的大小关系也较为模糊。未来研究可以考虑量化输入、输出的映射关系，更精确地探索不同映射关系对 AoA 效应的影响。第三，整合理论认为词汇习得年龄效应既来自表征的层面，也来自各表征间的联结。这一理论虽然整合了表征理论和任意映射理论的观点，但具体观点还相对概括，并未进一步澄清内在的影响机制，例如表征层和表征间联结这两种来源对 AoA 效应的贡献是否独立，即是否存在相互作用，以及两种来源的贡献大小关系等细节问题有待进一步研究(Chang & Lee, 2020; Elsherif et al., 2023; Wang et al., 2023)。AoA 效应的产生与词汇的获得与发展存在密切关联，对于揭示大脑的学习机制具有重要的理论意义。进一步研究可以通过计算模拟 AoA 在不同映射关系的语言加工过程中的表现，以及使用神经影像学方法进一步揭示脑在结构和功能上对映射关系的表现，以系统深入地揭示 AoA 效应的认知和神经机制，厘清整合理论的模糊之处。第四，本研究中主要关注了四类研究较多的任务，其他任务比如图画书写任务(Bonin et al., 2002)、延迟命名任务(Barry et al., 2001)、快速命名任务(Ghyselinck, Lewis, Brysbaert, 2004)，以及眼动研究(Carter & Luke, 2020)等并未纳入考虑。随着研究的积累，未来研究可以在现有分析技术上考虑纳入更多的任务类型以及技术手段，通过多任务多模态的结果综合论证 AoA 效应的认知机制，为当前的理论争论提供可信的数据支持。

参考文献（*元分析用到的参考文献）

*白利莉，陈宝国. (2011). 汉字习得的年龄对词类信息加工的影响. *心理科学, 34*(2), 343−347. https://doi.org/10.16719/j.cnki.1671-6981.2011.02.031

白学军，王丽红，吕勇，胡伟. (2007). 听觉呈现条件下词汇获得年龄效应的 ERP 研究. (摘要). *第三届全国*

脑与认知科学学术研讨会暨脑与认知科学国际研讨会(pp.24). 昆明云南.
*白学军, 王丽红, 吕勇, 胡伟. (2010). 词汇的获得年龄效应:ERP 研究. *心理学探新, 30*(1), 21−26. https://doi.org/10.3969/j.issn.1003-5184.2010.01.005
*陈宝国, 王立新, 王璐璐, 彭聃龄. (2004). 词汇习得年龄和频率对词汇识别的影响. *心理科学, 27*(5), 1060−1064. https://doi.org/10.16719/j.cnki.1671-6981.2004.05.009
*陈宝国, 尤文平, 周会霞. (2007). 汉语词汇习得的年龄效应:语义假设的证据. *心理学报, 39*(1), 9−17.
*陈宝国, 尤文平, 张亚峰, 刘文焕. (2010). 汉字早期字形加工阶段的习得年龄效应. *心理科学, 33*(3), 726−728.
*陈新葵, 张积家. (2010). 影响汉语动词、名词识别因素的回归分析. *心理科学, 33*(1), 60−63. doi:10.16719/j.cnki.1671-6981.2010.01.014.
*陈永香, 朱莉琪. (2015). 词汇命名中口语和书面语习得年龄效应的分离. (摘要). *中国心理学会发展心理专业委员会第十三届学术年会摘要集*(pp.72). 长春吉林.
郭英, 田鑫, 胡东, 白书琳, 周蜀溪. (2023). 羞愧对亲社会行为影响的三水平元分析. *心理科学进展, 31*(3), 371−385.
*郝美玲, 刘友谊, 舒华, 程辰曦. (2003). 汉语图画命名中获得年龄的作用. *心理与行为研究, 1*(4), 268−273.
*郝爽, 陈俊, 薛路芳. (2014). 听觉条件下词汇获得年龄效应产生机制的初探. *心理科学, 37*(2), 277−282. doi:10.16719/j.cnki.1671-6981.2014.02.006.
蓝媛美, 李超平, 王佳燕, 孟雪. (2022). 员工跨界行为的收益与代价:元分析的证据. *心理学报, 54*(6), 665−683.
*李丛, 张清芳, 黄韧. (2017). 词汇获得年龄在物体和动作图画命名中的不同作用. *心理学探新, 37*(3), 220−225.
*娄昊, 李丛, 张清芳. (2019). 习得年龄对客体和动作图画口语命名的不同影响:ERP 研究. *心理学报, 51*(2), 143−153. https://doi.org/10.3724/sp.J.1041.2019.00143
孟现鑫, 颜晨, 俞德霖, 高树玲, 傅小兰. (2024). 童年创伤与网络成瘾关系的三水平元分析. *心理科学进展, 32*(7), 1087−1103. https://doi.org/10.3724/SP.J.1042.2024.01087
孟现鑫, 俞德霖, 陈怡静, 张玲, 傅小兰. (2023). 儿童期创伤与共情的关系:一项三水平元分析. *心理学报, 55*(8), 1285−1300. https://doi.org/10.3724/SP.J.1041.2023.01285
彭聃龄, 郭德俊, 张素兰. (1985). 再认性同一判断中汉字信息的提取. *心理学报, 3*, 227−234.
*王丽红, 王永妍, 闫国利. (2010). 词汇获得年龄效应的眼动研究. *心理与行为研究, 8*(4), 289−295.
诸彦含, 贺彬, 孙蕾. (2024). 状态权力感对亲社会行为的影响:一项三水平元分析. *心理科学进展, 32*(11), 1786−1799.
张积家, 王娟, 印丛. (2014). 声符和义符在形声字语音、语义提取中的作用——来自部件知觉眼动研究的证据. *心理学报, 46*(7), 885−900. https://doi.org/10.3724/SP.J.1041.2014.00885
*张振军, 丁国盛, 陈宝国. (2011). 汉字习得的年龄效应:语音完整性假设的检验. *心理发展与教育, 27*(6), 577−583.
赵子卿, 余锦婷, 陈嘉彦, 王芸茹, 黄佳, 陈楚侨. (2024). 精神病临床高危人群的症状和功能改变:一项系统综述和三水平元分析. *心理科学进展, 33*(1), 42−61.
*Adorni, R., Manfredi, M., & Proverbio, A. M. (2013). Since when or how often? Dissociating the roles of age of acquisition (AoA) and lexical frequency in early visual word processing. *Brain and Language, 124*(1), 132−141. https://doi.org/10.1016/j.bandl.2012.11.005
*Alario, F. X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods Instruments & Computers, 36*(1), 140−155. https://doi.org/10.3758/bf03195559
Anderson, K., & Cottrell, G. (2001). Age of acquisition in connectionist networks. *Proceedings of the 23rd annual meeting of the cognitive science society*. Mahwah, NJ: Erlbaum.
Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *Tutorials in Quantitative Methods for Psychology, 12*(3), 154−174. https://doi.org/10.20982/tqmp.12.3.p154
*Baek, H., Gordon, P. C., & Choi, W. (2024). Effects of age and word frequency on Korean visual word recognition: Evidence from a web-based large-scale lexical-decision task. *Psychology and Aging, 39*(3), 231−244. https://doi.org/10.1037/pag0000793
*Bakhtiar, M., Nilipour, R., & Weekes, B. S. (2013). Predictors of timed picture naming in Persian. *Behavior Research Methods, 45*(3), 834−841. https://doi.org/10.3758/s13428-012-0298-6
*Bakhtiar, M., Su, I. F., Lee, H. K., & Weekes, B. S. (2016). Neural correlates of age of acquisition on visual word recognition in Persian. *Journal of Neurolinguistics, 39*, 1−9. https://doi.org/10.1016/j.jneuroling.2015.12.001
Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology. Human Perception and Performance, 10*(3), 340−357. https://doi.org/10.1037/0096-1523.10.3.340
*Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*(2), 283−316. https://doi.org/10.1037/0096-3445.133.2.283
*Bangalore, S., Robson, H., & Astell, A. J. (2022). Standardizing norms for 180 coloured Snodgrass and Vanderwart pictures in Kannada language. *Plos One, 17*(4), e0266359. https://doi.org/10.1371/journal.pone.0266359
*Barry, C., Hirsh, K. W., Johnston, R. A., & Williams, C. L. (2001). Age of acquisition, word frequency, and the

locus of repetition priming of picture naming. *Journal of Memory and Language, 44*(3), 350−375. https://doi.org/10.1006/jmla.2000.2743
*Barry, C., Johnston, R. A., & Wood, R. F. (2006). Effects of age of acquisition, age, and repetition priming on object naming. *Visual Cognition, 13*(7−8), 911−927. https://doi.org/10.1080/13506280544000101
*Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency and name agreement. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 50*(3), 560−585. https://doi.org/10.1080/10.1080/783663595
Belke, E., Brysbaert, M., Meyer, A. S., & Ghyselinck, M. (2005). Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. *Cognition, 96*(2), B45−B54. https://doi.org/10.1016/j.cognition.2004.11.006
Ben-Shachar, M., Lüdecke, D., & Makowski, D. (2020). Effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software, 5*(56), 12815. https://doi.org/10.21105/joss.02815
*Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language, 50*(4), 456−476. https://doi.org/10.1016/j.jml.2004.02.001
*Bonin, P., Boyer, B., Méot, A., Fayol, M., & Droit, S. (2004). Psycholinguistic norms for action photographs in French and their relationships with spoken and written latencies. *Behavior Research Methods, Instruments, and Computers, 36*(1), 127−139.
*Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology, 93*(1), 89−114. https://doi.org/10.1348/000712602162463
*Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, and Computers, 35*(1), 158−167. https://doi.org/10.3758/BF03195507
*Brown, G. D. A., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition, 15*(3), 208−216. https://doi.org/10.3758/BF03197718
Brysbaert, M., & Ellis, A. W. (2016). Aphasia and age of acquisition: are early-learned words more resilient? *Aphasiology, 30*(11), 1240−1263. https://doi.org/10.1080/02687038.2015.1106439
Brysbaert, M., Lagrou, E., & Stevens, M. (2017). Visual word recognition in a second language: A test of the lexical entrenchment hypothesis with lexical decision times. *Bilingualism, 20*(3), 530−548. https://doi.org/10.1017/S1366728916000353
Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition, 13*(7−8), 992−1011. https://doi.org/10.1080/13506280544000165
*Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica, 104*(2), 215−226. https://doi.org/10.1016/s0001-6918(00)00021-4
*Burani, C., Arduino, L. S., & Barca, L. (2007). Frequency, not age of acquisition, affects Italian word naming. *European Journal of Cognitive Psychology, 19*(6), 828−866. https://doi.org/10.1080/09541440600847946
*Busch, J. L., Haeussler, F. S., Domahs, F., Timmermann, L., Weber, I., & Oehrn, C. R. (2022). German normative data with naming latencies for 283 action pictures and 600 action verbs. *Behavior Research Methods, 54*(2), 649−662. https://doi.org/10.3758/s13428-021-01647-w
Cao, F., Peng, D., Liu, L., Jin, Z., Fan, N., Deng, Y., & Booth, J. R. (2009). Developmental differences of neurocognitive networks for phonological and semantic processing in Chinese word reading. *Human Brain Mapping*, *30*(3), 797−809. https://doi.org/10.1002/hbm.20546
Card, N. A. (2016). *Applied meta-analysis for social science research*. The Guilford Press.
Carroll, J. B., & White, M. N. (1973a). Age-of-acquisition norms for 220 picturable nouns. *Journal of Verbal Learning and Verbal Behavior, 12*(5), 563−576. https://doi.org/https://doi.org/10.1016/S0022-5371(73)80036-2
*Carroll, J. B., & White, M. N. (1973b). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology, 25*(1), 85−95. doi:10.1080/14640747308400325
Carter, B. T., & Luke, S. G. (2020). Best practices in eye tracking research. *International Journal of Psychophysiology*, *155*, 49−62. https://doi.org/10.1016/j.ijpsycho.2020.05.010
*Catling, J. C., Dent, K., & Williamson, S. (2008). Age of acquisition, not word frequency affects object recognition: Evidence from the effects of visual degradation. *Acta Psychologica, 129*(1), 130−137. https://doi.org/10.1016/j.actpsy.2008.05.005
*Catling, J. C., & Elsherif, M. M. (2020). The hunt for the age of acquisition effect: It's in the links. *Acta Psychologica, 209* , 103138. https://doi.org/10.1016/j.actpsy.2020.103138
*Catling, J. C., & Johnston, R. A. (2006a). Age of acquisition effects on an object-name verification task. *British Journal of Psychology, 97*(1), 1−18. https://doi.org/10.1348/000712605x53515
*Catling, J. C., & Johnston, R. A. (2006b). Effects of age of acquisition and priming on picture naming. *Quarterly Journal of Experimental Psychology, 59*(8), 1443−1453. https://doi.org/10.1080/17470210500214291
*Catling, J. C., & Johnston, R. A. (2006c). The effects of age of acquisition on an object classification task. *Visual Cognition, 13*(7−8), 968−980. https://doi.org/10.1080/13506280544000138
*Catling, J. C., & Johnston, R. A. (2009). The varying effects of age of acquisition. *Quarterly Journal of Experimental Psychology, 62*(1), 50−62. https://doi.org/10.1080/17470210701814352
*Chalard, M., & Bonin, P. (2006). Age-of-acquisition effects in picture naming: Are they structural and/or semantic in nature? *Visual Cognition, 13*(7−8), 864−883. https://doi.org/10.1080/13506280544000084

Chang, Y. N., Monaghan, P., & Welbourne, S. (2019). A computational model of reading across development: Effects of literacy onset on language processing. *Journal of Memory and Language, 108*, 104025. https://doi .org/10.1016/j.jml.2019.05.003
*Chang, Y. N., & Lee, C. Y. (2020). Age of acquisition effects on traditional Chinese character naming and lexical decision. *Psychonomic Bulletin & Review, 27*(6), 1317−1324. https://doi.org/10.3758/s13423-020-01787-8
*Chen, B. G., Dent, K., You, W., & Wu, G. (2009). Age of acquisition affects early orthographic processing during Chinese character recognition. *Acta Psychologica, 130*(3), 196−203. https://doi.org/10.1016/j.actpsy.2008.12.004
*Chen, B. G., Zhou, H. X., Dunlap, S., & Perfetti, C. A. (2007). Age of acquisition effects in reading Chinese: Evidence in favour of the arbitrary mapping hypothesis. *British Journal of Psychology, 98*(3), 499−516. https://doi.org/10.1348/000712606x165484
Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods, 19*(2), 211−229. https://doi.org/10.1037/a0032968
Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155−159. https://doi.org/10.1037/0033-2909.112.1.155
*Colombo, L., & Burani, C. (2002). The influence of age of acquisition, root frequency, and context availability in processing nouns and verbs. *Brain and Language, 81*(1−3), 398−411. https://doi.org/10.1006/brln.2001.2533
*Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *Quarterly Journal of Experimental Psychology, 60*(8), 1072−1082. https://doi.org/10.1080/17470210701315467
Cortese, M. J., Toppi, S., Khanna, M. M., & Santo, J. B. (2020). AoA effects in reading aloud and lexical decision: Locating the (semantic) locus in terms of the number of backward semantic associations. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 73*(11), 2036−2044. https://doi.org/10.1177/1747021820940302
Cortese, M. J., Yates, M., Schock, J., & Vilks, L. (2018). Examining word processing via a megastudy of conditional reading aloud. *Quarterly Journal of Experimental Psychology, 71*(11), 2295−2313. https://doi.org/10.1177/1747021817741269
*Cuetos, F., & Alija, M. (2003). Normative data and naming times for action pictures. *Behavior Research Methods, Instruments, and Computers, 35*(1), 168−177. https://doi.org/10.3758/BF03195508
*Cuetos, F., & Barbon, A. (2006). Word naming in Spanish. *European Journal of Cognitive Psychology, 18*(3), 415−436. https://doi.org/10.1080/13594320500165896
*Cuetos, F., Barbón, A., Urrutia, M., & Domínguez, A. (2009). Determining the time course of lexical frequency and age of acquisition using ERP. *Clinical Neurophysiology, 120*(2), 285−294. https://doi.org/10.1016/j.clinph.2008.11.003
*Cuetos, F., Ellis, A. W., & Alvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, and Computers, 31*(4), 650−658. https://doi.org/10.3758/BF03200741
*Davies, R., Wilson, M., Cuetos, F., & Burani, C. (2014). Reading in Spanish and Italian: Effects of age of acquisition in transparent orthographies? *Quarterly Journal of Experimental Psychology, 67*(9), 1808−1825. https://doi.org/10.1080/17470218.2013.872155
*De Deyne, S., & Storms, G. (2007). Age-of-acquisition differences in young and older adults affect latencies in lexical decision and semantic categorization. *Acta Psychologica, 124*(3), 274−295. https://doi.org/10.1016/j.actpsy.2006.03.007
*Dell'Acqua, R., Lotto, L., & Remo, J. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, and Computers, 32*(4), 588−615. https://doi.org/10.3758/BF03200832
*den Hollander, J., Jonkers, R., Mariën, P., & Bastiaanse, R. (2019). Identifying the speech production stages in early and late adulthood by using electroencephalography. *Frontiers in Human Neuroscience, 13*, 298−298. https://doi.org/10.3389/fnhum.2019.00298
*Dent, K., Catling, J. C., & Johnston, R. A. (2007). Age of acquisition affects object recognition: Evidence from visual duration thresholds. *Acta Psychologica, 125*(3), 301−318. https://doi.org/10.1016/j.actpsy.2006.08.004
*Dewhurst, S. A., & Barry, C. (2006). Dissociating word frequency and age of acquisition: The Klein effect revived (and reversed). *Journal of Experimental Psychology-Learning Memory and Cognition, 32*(4), 919−924. https://doi.org/10.1037/0278-7393.32.4.919
Dirix, N., & Duyck, W. (2017). An eye movement corpus study of the age-of-acquisition effect. *Psychonomic Bulletin & Review, 24*(6), 1915−1921. https://doi.org/https://doi.org/10.3758/s13423-017-1233-8.
Dodell-Feder, D., & Tamir, D. I. (2018). Fiction reading hasa small positive impact on social cognition: A meta-analysis. *Journal of Experimental Psychology: General, 147*(11), 1713−1727.
Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455−463. https://doi.org/10.1111/j.0006-341X.2000.00455.x
*Egbert, M. H. Assink, Van Well, S., & Paul P. N. A. Knuijt. (2003). Contrasting effects of age of acquisition in lexical decision and letter detection. *The American Journal of Psychology, 116*(3), 367−387. https://doi.org/10.2307/1423499
Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629−634. https://doi.org/10.1136/bmj.315.7109.629

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning Memory and Cognition, 26*(5), 1103−1123. https://doi.org/10.1037/0278-7393.26.5.1103
*Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of experimental psychology. Learning, memory, and cognition, 24*(2), 515−523. https://doi.org/10.1037//0278-7393.24.2.515
Elsherif, M. M., & Catling, J. C. (2022). Age of acquisition effects on the decomposition of compound words. *Journal of Cognitive Psychology, 34*(3), 325−338. https://doi.org/10.1080/20445911.2021.2013246
*Elsherif, M. M., Catling, J. C., & Frisson, S. (2020). Two words as one: A multi-naming investigation of the age-of-acquisition effect in compound-word processing. *Memory & Cognition, 48*(4), 511−525. https://doi.org/10.3758/s13421-019-00986-6
Elsherif, M. M., Preece, E., & Catling, J. C. (2023). Age-of-acquisition effects: A literature review. *Journal of Experimental Psychology-Learning Memory and Cognition, 49*(5), 812−847. https://doi.org/10.1037/xlm0001215
*Fargier, R., & Laganaro, M. (2020). Neural dynamics of the production of newly acquired words relative to well-known words. *Brain Research, 1727*, 146557. https://doi.org/10.1016/j.brainres.2019.146557
*Fiebach, C. J., Friederici, A. D., Muller, K., von Cramon, D. Y., & Hernandez, A. E. (2003). Distinct brain representations for early and late learned words. *Neuroimage, 19*(4), 1627−1637. https://doi.org/10.1016/s1053-8119(03)00227-1
*Gerhand, S., & Barry, C. (1998). Word frequency effects in oral reading are not merely age-of-acquisition effects in disguise. *Journal of Experimental Psychology-Learning Memory and Cognition, 24*(2), 267−283. https://doi.org/10.1037/0278-7393.24.2.267
*Gerhand, S., & Barry, C. (1999a). Age-of-acquisition and frequency effects in speeded word naming. *Cognition, 73*(2), B27−B36. https://doi.org/https://doi.org/10.1016/S0010-0277(99)00052-9
*Gerhand, S., & Barry, C. (1999b). Age of acquisition, word frequency, and the role of phonology in the lexical decision task. *Memory & Cognition, 27*(4), 592−602. https://doi.org/10.3758/bf03211553
*Gilhooly, K. J. (1984). Word age-of-acquisition and residence time in lexical memory as factors in word naming. *Current Psychological Research & Reviews, 3*(2), 24−31. https://doi.org/10.1007/BF02686547
*Gilhooly, K. J., & Gilhooly, M. L. (1979). Age-of-acquisition effects in lexical and episodic memory tasks. *Memory & Cognition, 7*(3), 214−223. https://doi.org/10.3758/BF03197541
*Gilhooly, K. J., & Logie, R. H. (1981a). Word age-of-acquisition, reading latencies and auditory recognition. *Current Psychology, 1*(3−4), 251−262. https://doi.org/10.1007/BF03186735
*Gilhooly, K. J., & Logie, R. H. (1981b). Word age-of-acquisition and visual recognition thresholds. *Current Psychology, 1*(3−4), 215−225. https://doi.org/10.1007/BF03186732
*Gilhooly, K. J., & Logie, R. H. (1982). Word age-of-acquisition and lexical decision making. *Acta Psychologica, 50*(1), 21−34. doi:10.1016/0001-6918(82)90048-8
*Ghyselinck, M., Custers, R., & Brysbaert, M. (2004). The effect of age of acquisition in visual word processing: Further evidence for the semantic hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 30*(2), 550−554. https://doi.org/10.1037/0278-7393.30.2.550
*Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica, 115*(1), 43−67. https://doi.org/10.1016/j.actpsy.2003.11.002
*González-Nosti, M., Barbón, A., Rodríguez-Ferreiro, J., & Cuetos, F. (2014). Effects of the psycholinguistic variables on the lexical decision task in Spanish: A study with 2,765 words. *Behavior Research Methods, 46*(2), 517−525. https://doi.org/10.3758/s13428-013-0383-5
Grainger, J., & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Language and Linguistics Compass*, *3*(1), 128−156. https://doi.org/10.1111/j.1749-818X.2008.00121.x
Haman, E., Luniewska, M., Hansen, P., Simonsen, H. G., Chiat, S., Bjekic, J., … Armon-Lotem, S. (2017). Noun and verb knowledge in monolingual preschool children across 17 languages: Data from Cross-linguistic Lexical Tasks (LITMUS-CLT). *Clinical Linguistics & Phonetics, 31*(11−12), 818−843. https://doi.org/10.1080/02699206.2017.1308553
Han, J. I., & Kim, J. Y. (2017). The influence of orthography on the production of alphabetic, second-language allophones by speakers of a non-alphabetic language. *Journal of Psycholinguistic Research, 46*(4), 963−982. https://doi.org/10.1007/s10936-016-9474-7
Han, J. I., Kim, J. Y., & Choi, T. H. (2021). The role of orthography in lexical processing of the phonological variants in second language. *Journal of Psycholinguistic Research, 50*(2), 437−445. https://doi.org/10.1007/s10936-020-09725-4
*Havelka, J., & Tomita, I. (2006). Age of acquisition in naming Japanese words. *Visual Cognition, 13*(7−8), 981−991. https://doi.org/10.1080/13506280544000156
*Hernandez, A. E., & Fiebach, C. J. (2006). The brain bases of reading late learned words: Evidence from functional MRI. *Visual Cognition, 13*(7−8), 1027−1043. https://doi.org/10.1080/13506280544000183
Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin*, *133*(4), 638−650. https://doi.org/10.1037/0033-2909.133.4.638
Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557−560.
*Holmes, S. J., & Ellis, A. W. (2006). Age of acquisition and typicality effects in three object processing tasks.

Visual Cognition, 13(7−8), 884−910. https://doi.org/10.1080/13506280544000093
*Izura, C., & Hernández-Muñoz, N. (2017). The role of semantics in Spanish word recognition: An insight from lexical decision and categorization tasks. *Open Linguistics, 3*(1), 500−515. https://doi.org/10.1515/opli-2017-0025
*Izura, C., & Playfoot, D. (2012). A normative study of acronyms and acronym naming. *Behavior Research Methods, 44*(3), 862−889. https://doi.org/10.3758/s13428-011-0175-8
Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131*(5), 684−712. https://doi.org/10 .1037/0033-2909.131.5.684
Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition, 13*(7−8), 789−845. https://doi.org/10.1080/13506280544000066
*Johnson, Carla J. & Clark, James M. (1988). Children's picture naming difficulty and errors: Effects of age of acquisition, uncertainty, and name generality. *Applied Psycholinguistics, 9*(4), 351−365. doi:10.1017/S0142716400008055
*Johnston, R. A., Dent, K., Humphreys, G. W., & Barry, C. (2010). British-English norms and naming times for a set of 539 pictures: The role of age of acquisition. *Behavior Research Methods, 42*(2), 461−469. https://doi.org/10.3758/BRM.42.2.461
*Karimi, H., & Diaz, M. (2020). When phonological neighborhood density both facilitates and impedes: Age of acquisition and name agreement interact with phonological neighborhood during word production. *Memory & Cognition, 48*(6), 1061−1072. https://doi.org/10.3758/s13421-020-01042-4
Kauschke, C., & Von Frankenberg, J. (2008). The differential influence of lexical parameters on naming latencies in German. A study on noun and verb picture naming. *Journal of Psycholinguistic Research, 37*(4), 243−257. https://doi.org/10.1007/s10936-007-9068-5
*Khwaileh, T., Body, R., & Herbert, R. (2014). A normative database and determinants of lexical retrieval for 186 Arabic nouns: Effects of psycholinguistic and morpho-syntactic variables on naming latency. *Journal of Psycholinguistic Research, 43*(6), 749−769. https://doi.org/10.1007/s10936-013-9277-z
Kuperman, V. (2013). Accentuate the positive: Semantic access in English compounds. *Frontiers in Psychology, 4*, 203−203. https://doi.org/10.3389/fpsyg.2013.00203
*Laganaro, M. (2014). ERP topographic analyses from concept to articulation in word production studies. *Frontiers in Psychology, 5*, 493−493. https://doi.org/10.3389/fpsyg.2014.00493
*Laganaro, M., & Perret, C. (2011). Comparing electrophysiological correlates of word production in immediate and delayed naming through the analysis of word age of acquisition effects. *Brain Topography, 24*(1), 19−29. https://doi.org/10.1007/s10548-010-0162-x
*Laganaro, M., Valente, A., & Perret, C. (2012). Time course of word production in fast and slow speakers: A high density ERP topographic study. *Neuroimage, 59*(4), 3881−3888. https://doi.org/10.1016/j.neuroimage.2011.10.082
*Lambon Ralph, M. A., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence: Empirical and computational evidence. *Visual Cognition, 13*(7−8), 928−948. https://doi.org/10.1080/13506280544000110
*Laws, K. R., Leeson, V. C., & Gale, T. M. (2002). The effect of 'masking' on picture naming. *Cortex, 38*(2), 137−147. https://doi.org/10.1016/s0010-9452(08)70646-4
Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1−75. https://doi.org/10.1017/s0140525x99001776
Levitt, A. G., & Healy, A. F. (1985). The roles of phoneme frequency, similarity, and availability in the experimental elicitation of speech errors. *Journal of Memory and Language, 24*(6), 717−733. https://doi.org/https://doi.org/10.1016/0749-596X(85)90055-5
Light, R. J., & Pillemer, D. B. (1984). *Summing up the science of reviewing research*. Harvard University Press. https://doi.org/10.2307/j.ctvk12px9
Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Sage Publications, Inc.
*Liu, Y., Hao, M., Li, P., & Shu, H. (2011). Timed picture naming norms for Mandarin Chinese. *Plos One, 6*(1), e16505. https://doi.org/10.1371/journal.pone.0016505
*Liu, Y., Hao, M., Shu, H., Tan, L. H., & Weekes, B. S. (2008). Age-of-acquisition effects on oral reading in Chinese. *Psychonomic Bulletin & Review, 15*(2), 344−350. https://doi.org/10.3758/pbr.15.2.344
Lukatela, G., & Turvey, M. T. (1994). Visual lexical access is initially phonological: I. Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology-General, 123*(2), 107−128. https://doi.org/10.1037/0096-3445.123.2.107
*Menenti, L., & Burani, C. (2007). What causes the effect of age of acquisition in lexical processing? *Quarterly Journal of Experimental Psychology, 60*(5), 652−660. https://doi.org/10.1080/17470210601100126
*Meschyan, G., & Hernandez, A. (2002). Age of acquisition and word frequency: Determinants of object-naming speed and accuracy. *Memory & Cognition, 30*(2), 262−269. https://doi.org/10.3758/bf03195287
*Monaghan, J., & Ellis, A. W. (2002a). Age of acquisition and the completeness of phonological representations. *Reading and Writing, 15*(7−8), 759−788.
*Monaghan, J., & Ellis, A. W. (2002b). What exactly interacts with spelling-sound consistency in word naming? *Journal of Experimental Psychology-Learning Memory and Cognition, 28*(1), 183−206. https://doi.org/10.1037//0278-7393.28.1.183
Monaghan, P., & Ellis, A. W. (2010). Modeling reading development: Cumulative, incremental learning in a computational model of word naming. *Journal of Memory and Language, 63*(4), 506−525. https://doi.org/https://doi.org/10.1016/j.jml.2010.08.003

Moore, V., Smith-Spark, J., & Valentine, T. (2004). The effects of age of acquisition on object perception. *European Journal of Cognitive Psychology, 16*(3), 417−439. https://doi.org/10.1080/09541440340000097
Morris, P. E. (1981). Age of acquisition, imagery, recall, and the limitations of multiple-regression analysis. *Memory & Cognition, 9*(3), 277−282. https://doi.org/10.3758/bf03196961
Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 50*(3), 528−559. https://doi.org/10.1080/027249897392017
*Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 116−133. https://doi.org/10.1037/0278-7393.21.1.116
*Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology, 91*(2), 167−180. https://doi.org/10.1348/000712600161763
*Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition, 20*(6), 705−714. https://doi.org/10.3758/BF03202720
*Morrison, C. M., & Gibbons, Z. C. (2006). Lexical determinants of semantic processing speed. *Visual Cognition, 13*(7−8), 949−967. https://doi.org/10.1080/13506280544000129
*Morrison, C. M., Hirsh, K. W., Chappell, T., & Ellis, A. W. (2002). Age and age of acquisition: An evaluation of the cumulative frequency hypothesis. *European Journal of Cognitive Psychology, 14*(4), 435−459. https://doi.org/10.1080/09541440143000159
*Morrison, C. M., Hirsh, K. W., & Duggan, G. B. (2003). Age of acquisition, ageing, and verb production: Normative and experimental data. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 56*(4), 705−730. https://doi.org/10.1080/02724980244000594
*Nagy, W., Anderson, R. C., Schommer, M., Scott, J. A., & Stallman, A. C. (1989). Morphological families in the internal lexicon. *Reading Research Quarterly, 24*(3), 262−282. https://doi.org/10.2307/747770
*Navarrete, E., Scaltritti, M., Mulatti, C., & Peressotti, F. (2013). Age-of-acquisition effects in delayed picture-naming tasks. *Psychonomic Bulletin and Review, 20*(1), 148−153. https://doi.org/10.3758/s13423-012-0310-2
National Institutes of Health. (2018). *Quality assessment tool for observationalcohort and crosssectional studies.* Retrieved July, 2021, from https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools
*Nishimoto, T., Miyawaki, K., Ueda, T., Une, Y., & Takahashi, M. (2005). Japanese normative set of 359 pictures. *Behavior Research Methods, 37*(3), 398−416. https://doi.org/10.3758/BF03192709
Pastor, D. A., & Lazowski, R. A. (2018). On the multilevel nature of meta-analysis: A tutorial, comparison of software programs, and discussion of analytic choices. *Multivariate Behavioral Research, 53*(1), 74−89. https://doi.org/10.1080/00273171.2017.1365684
*Perret, C., Bonin, P., & Laganaro, M. (2014). Exploring the multiple-level hypothesis of AoA effects in spoken and written object naming using a topographic ERP analysis. *Brain and Language, 135*, 20−31. https://doi.org/10.1016/j.bandl.2014.04.006
*Pind, J., & Tryggvadóttir, H. B. (2002). Determinants of picture naming times in Icelandic. *Scandinavian Journal of Psychology, 43*(3), 221−226. https://doi.org/10.1111/1467-9450.00290
*Ploetz, D. M., & Yates, M. (2016). Age of acquisition and imageability: A cross-task comparison. *Journal of Research in Reading, 39*(1), 37−49. https://doi.org/10.1111/1467-9817.12040
*Räling, R., Holzgrefe-Lang, J., Schröder, A., & Wartenburger, I. (2015). On the influence of typicality and age of acquisition on semantic processing: Diverging evidence from behavioural and ERP responses. *Neuropsychologia, 75*, 186−200. https://doi.org/10.1016/j.neuropsychologia.2015.05.031
*Räling, R., Hanne, S., Schröder, A., Keßler, C., & Wartenburger, I. (2017). Judging the animacy of words: The influence of typicality and age of acquisition in a semantic decision task. *Quarterly Journal of Experimental Psychology*, *70*(10), 2094−2104. https://doi.org/10.1080/17470218.2016.1223704
Raman, I. (2006). On the age-of-acquisition effects in word naming and orthographic transparency: Mapping specific or universal? *Visual Cognition, 13*(7−8), 1044−1053. https://doi.org/10.1080/13506280500153200
*Raman, I. (2011). The role of age of acquisition in picture and word naming in dyslexic adults. *British Journal of Psychology, 102*, 328−339. https://doi.org/10.1348/000712610x522572
*Raman, I. (2018). The role of context on age of acquisition effect: Strategic control in word naming in Turkish. In *Psycholinguistics and cognition in language processing* (pp. 19-48). IGI Global.
*Raman, I., Raman, E. İkier S., Kilecioğlu, E., Uzun, E. D., & Zeyveli, Ş. (2018). Differential effects of age of acquisition and frequency on memory: Evidence from free recall of pictures and words in Turkish. *Writing Systems Research, 10*, 1−14.
Rochford, G., & Williams, M. (1962a). I The relationship between nominal dysphasia and the acquisition of vocabulary in childhood. *Journal of Neurology, Neurosurgery & Psychiatry, 25*(3), 222−227. https://doi.org/10.1136/jnnp.25.3.222
Rochford, G., & Williams, M. (1962b). II Experimental production of naming disorders in normal people. *Journal of Neurology, Neurosurgery & Psychiatry, 25*(3), 228−233. https://doi.org/10.1136/jnnp.25.3.228
*Roodenrys, S., Hulme, C., Alban, J., Ellis, A. W., & Brown, G. D. A. (1994). Effects of word frequency and age of acquisition on short-term memory span. *Memory & Cognition, 22*(6), 695−701. https://doi.org/10.3758/BF03209254
Rothstein, H., Sutton, A. J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments.* Wiley.
Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word

recognition. *Journal of Verbal Learning and Verbal Behavior, 10*(6), 645−657. https://doi.org/https://doi.org/10.1016/S0022-5371(71)80071-3
*Spataro, P., Lpngobardi, E., Saraulli, D., & Rossi-Arnaud, C. (2013). Interactive effects of age-of-acquisition and repetition priming in the lexical decision task: A multiple-loci account. *Experimental Psychology, 60*(4), 235−242. https://doi.org/10.1027/1618-3169/a000192
*Schwier, V., Boyer, B., Meot, A., Bonin, P., & Laganaro, M. (2004). French normative data and naming times for action pictures. *Behavior Research Methods Instruments & Computers, 36*(3), 564−576. https://doi.org/10.3758/bf03195603
Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental of word recognition and naming. *Psychological Review, 96*(4), 523−568. https://doi.org/10.1037/0033-295x.96.4.523
*Sereno, S. C., & O'Donnell, P. J. (2009). Participant and word gender in age of acquisition effects: The role of gender socialization. *Sex Roles, 61*(7−8), 510−518. https://doi.org/10.1007/s11199-009-9649-x
*Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica, 119*(2), 159−187. https://doi.org/10.1016/j.actpsy.2005.01.002
*Shao, Z., Roelofs, A., & Meyer, A. S. (2014). Predicting naming latencies for action pictures: Dutch norms. *Behavior Research Methods, 46*(1), 274−283. https://doi.org/10.3758/s13428-013-0358-6
*Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, and Computers, 28*(4), 516−536. https://doi.org/10.3758/BF03200540
*Song, D., & Li, D. (2021). Psycholinguistic norms for 3,783 two-character words in simplified Chinese. *Sage Open, 11*(4). https://doi.org/10.1177/21582440211054495
Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29*(1), 41−78. https://doi.org/10.1207/s15516709cog2901_3
*Tainturier, M. J., Tamminen, J., & Thierry, G. (2005). Age of acquisition modulates the amplitude of the P300 component in spoken word recognition. *Neuroscience Letters, 379*(1), 17−22. https://doi.org/10.1016/j.neulet.2004.12.038
*Turner, J. E., Valentine, T., & Ellis, A. W. (1998). Contrasting effects of age of acquisition and word frequency on auditory and visual lexical decision . *Memory and Cognition, 26*(6), 1282−1291. https://doi.org/10.3758/BF03201200
*Urooj, U., Cornelissen, P. L., Simpson, M. I. G., Wheat, K. L., Woods, W., Barca, L., & Ellis, A. W. (2014). Interactions between visual and semantic processing during object recognition revealed by modulatory effects of age of acquisition. *NeuroImage, 87*, 252−264. https://doi.org/10.1016/j.neuroimage.2013.10.058
*Valente, A., Buerki, A., & Laganaro, M. (2014). ERP correlates of word production predictors in picture naming: A trial by trial multiple regression analysis from stimulus onset to response. *Frontiers in Neuroscience, 8*, 390. https://doi.org/10.3389/fnins.2014.00390
*Vitkovitch, M., & Tyrrell, L. (1995). Sources of disagreement in object naming. *The Quarterly Journal of Experimental Psychology Section A, 48*(4), 822−848. https://doi.org/10.1080/14640749508401419
Wang, J., Jiang, X., & Chen, B. (2023). Second language age of acquisition effects in a word naming task: A regression analysis of ERP data. *Journal of Neurolinguistics, 66*, 101125. https://doi.org/10.1016/j.jneuroling.2023.101125
*Wang, Z., Zhang, L., & Xuan, B. (2024). Age-related Chinese word recognition across different AoA and parts of speech. *Current Psychology, 43*(22), 19939−19952. https://doi.org/10.1007/s12144-024-05794-z
Weekes, B. (2011). Age of acquisition effects on Chinese character recognition: Evidence from EEG. *Procedia - Social and Behavioral Sciences, 23*, 67−68. https://doi.org/https://doi.org/10.1016/j.sbspro.2011.09.173
Weekes, B. S., Chan, A., Kwok, J. S. W., Hia Tan, L., & Jin, Z. (2004). AoA effects on Chinese language processing: An fMRI study. *Brain and Language, 91*(1), 33−34. https://doi.org/10.1016/j.bandl.2004.06.020
*Weekes, B. S., Chan, A., & Tan, L. H. (2008). Effects of age of acquisition on brain activation during Chinese character recognition. *Neuropsychologia, 46*(7), 2086−2090. https://doi.org/10.1016/j.neuropsychologia.2008.01.020
*Weekes, B. S., Shu, H., Hao, M., Liu, Y., & Tan, L. H. (2007). Predictors of timed picture naming in Chinese. *Behavior Research Methods, 39*(2), 335−342. https://doi.org/10.3758/bf03193165
*Wilson, M. A., Cuetos, F., Davies, R., & Burani, C. (2013). Revisiting age-of-acquisition effects in Spanish visual word recognition: The role of item imageability. *Journal of Experimental Psychology-Learning Memory and Cognition, 39*(6), 1842−1859. https://doi.org/10.1037/a0033090
*Wilson, M. A., Ellis, A. W., & Burani, C. (2012). Age-of-acquisition affects word naming in Italian only when stress is irregular. *Acta Psychologica, 139*(3), 417−424. https://doi.org/10.1016/j.actpsy.2011.12.012
*Wolna, A., Luniewska, M., Haman, E., & Wodniecka, Z. (2023). Polish norms for a set of colored drawings of 168 objects and 146 actions with predictors of naming performance. *Behavior Research Methods, 55*(5), 2706−2732. https://doi.org/10.3758/s13428-022-01923-3
*Xu, X., Li, J., & Guo, S. (2021). Age of acquisition ratings for 19,716 simplified Chinese words. *Behavior Research Methods, 53*(2), 558−573. https://doi.org/10.3758/s13428-020-01455-8
*Yamazaki, M., Ellis, A. W., Morrison, C. M., & Ralph, M. A. L. (1997). Two age of acquisition effects in the reading of Japanese Kanji. *British Journal of Psychology, 88*(3), 407−421.https://doi.org/10.1111/j.2044-8295.1997.tb02648.x
*Young, A. W., Bion, P. J., Ellis, A. W. (1982). Age of reading acquisition does not affect visual hemifield asymmetries for naming imageable nouns. *Cortex, 18*(3), 477−482.
*Yum, Y. N., & Law, S. P. (2019). Interactions of age of acquisition and lexical frequency effects with phonological regularity: An ERP study. *Psychophysiology, 56*(10), e13433.

https://doi.org/10.1111/psyp.13433
Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language, 47*(1), 1−29. https://doi.org/10.1006/jmla.2001.2834
Zevin, J. D., & Seidenberg, M. S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory & Cognition, 32*(1), 31−38. https://doi.org/10.3758/bf03195818
*Zhang, M., Liu, Z., Botezatu, M. R., Dang, Q., Yuan, Q., Han, J., … Guo, T. (2023). A large-scale database of Chinese characters and words collected from elementary school textbooks. *Behavior Research Methods, 56*(5), 4732−4757. https://doi.org/10.3758/s13428-023-02214-1
*Zhang, Q., Zhou, Y., & Lou, H. (2022). The dissociation between age of acquisition and word frequency effects in Chinese spoken picture naming. *Psychological Research-Psychologische Forschung, 86*(6), 1918−1929. https://doi.org/10.1007/s00426-021-01616-0

Cognitive mechanisms underlying the age of acquisition effects: Insights from a three-level meta-analysis

ZHANG Lu, WANG Ziqian, Zhang Qingfang

(*Department of Psychology, Renmin University of China, Beijing* 100872*, China*)

Abstract: Age of acquisition (AoA) refers to the age at which an individual first understands the meaning of a word. Researches revealed that the processing speed of early acquired words is faster and the accuracy is higher, called the AoA effect. To determine the cognitive mechanism underlying the AoA effect, researchers have proposed The Multiple Loci Account, Representation Theory, and Arbitrary Mapping Hypothesis. In order to systematically investigate the cognitive mechanism of the AoA effect, the present study employed a three-level meta-analysis to quantitatively analyze existing literature on the AoA effect. Through the retrieval of articles published before May 15, 2024, the current meta-analysis identified 130 studies, with 10697 participants and 271 effect sizes. Analysis revealed that the effect size of the AoA effect was Hedge's $g = 0.47$($p < 0.001$, 95%CI = [0.41, 0.52]). Additionally, the effect size was moderated by task type and word frequency, mainly depending on the semantic relevance across different tasks. However, no significant moderating effects were observed for the target cognitive process, writing system, or the phonological relevance across different tasks. The results of the current meta-analysis indicated that AoA mainly affects the semantic processing stage, and primarily supports the Representation Theory, with early-acquired words having a richer semantic network, thereby facilitating the word processing process.

Keywords: picture naming task, lexical decision task, word naming task, age of acquisition effects, three-level meta-analysis